

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 10, 2005, 19:08:10 ; Search time 42 Seconds  
(without alignments)  
1459.288 Million cell updates/sec

Title: US-10-623-429-9  
Perfect score: 3385  
Sequence: 1 MASAEMRERLEAPDRAPV.....VNVDTARAADLFVSMWMSR 637

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues  
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 79: \*  
1: pir1: \*  
2: pir2: \*  
3: pir3: \*  
4: pir4: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2333	68.9	635	1 WMBE6	capsid protein - h
2	1262	37.3	306	2 A54862	HSV-1 proteinase -
3	1079.5	31.9	646	1 WZBSC8	68.6K capsid prote
4	1056	31.2	647	2 T42579	capsid protein 35
5	943	27.9	522	2 S52216	viral proteinase -
6	824.5	24.4	605	1 WZB333	gene 33 protein -
7	401	11.8	586	1 A43675	capsid protein p40
8	383.5	11.3	643	2 S5510	polyprotein - equi
9	372.5	11.0	605	1 QQB33R	BVRP2 (SC-RP3) pro
10	331	9.8	459	2 T42930	hypothetical prote
11	322	9.5	590	1 WMBSC8	64K capsid assembl
12	317.5	9.4	708	1 QQB33B	UL80 protein - hum
13	301.5	8.9	475	1 WZBEM4	gene 17 protein -
14	280	8.3	839	2 T04859	extensin homolog F
15	278	8.2	524	2 T03112	minor capsid scaff
16	278	8.2	697	2 T42615	fused proteinase-c
17	272.5	8.1	528	2 T44013	proteinase / in-fr
18	270.5	8.0	528	2 T44200	probable proteinase
19	269	7.9	620	2 S06733	hydroxyproline-ric
20	264	7.8	2715	2 T13049	eyelid - fruit fly
21	262.5	7.8	519	2 S55470	proteinase - human
22	258.5	7.6	760	2 T06291	extensin homolog T
23	254.5	7.5	574	2 T43556	Wiskott-Aldrich sy
24	254.5	7.5	574	2 T38819	wiskott-aldrich sy
25	253.5	7.5	1184	2 G01763	atrophin-1 - human
26	251	7.4	464	2 S22597	extensin - Volvox
27	250	7.4	1494	2 T14355	protein-tyrosine-p
28	247.5	7.3	1184	2 S50832	atrophin-1 - human
29	246.5	7.3	1188	2 S49915	extensin-like prot

hypothetical prote  
hypothetical prote  
AAA protein I4171.  
hypothetical prote  
99.7K hypothetical  
BHLFI protein - hu  
extensin precursor  
extensin homolog F  
hypothetical prote  
probable proteinase  
unknown protein li  
extensin - almond  
hypothetical prote  
proline-rich prot  
extensin-like prot  
UL36 protein - hum

## ALIGNMENTS

### RESULT 1

WMBE6  
capsid protein - human herpesvirus 1 (strain 17)  
C;Species: human herpesvirus 1  
C;Date: 31-Dec-1989 #sequence\_revision 31-Dec-1989 #text\_change 09-Jul-2004  
C;Accession: H30084

R;McGeoch, D.J.; Dalrymple, M.A.; Davison, A.J.; Dolan, A.; Frame, M.C.; McNab, D.; Perr  
J. Gen. Virol. 69, 1531-1574, 1988

A;Title: The complete DNA sequence of the long unique region in the genome of herpes sim  
A;Reference number: A30083; MUID:88274327; PMID:2839594  
A;Accession: H30084

A;Status: nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-635 <WCG>

A;Cross-references: UNIPROT:P10210; GB:X14112; NID:G1944536; PIDN:CAA32318.1; PID:G59526  
C;Genetics:  
C;Superfamily: varicella-zoster virus gene 33 protein

A;Gene: UL26  
C;Keywords: capsid protein

Query Match 68.9%; Score 2333; DB 1; Length 635;  
Best Local Similarity 74.0%; Pred. No. 6.6e-114;  
Matches 477; Conservative 30; Mismatches 120; Indels 18; Gaps 7;

Qy 1 MASAEMRERLEAPDRAPVIYVAGFLALYDSGDELALDPDPTVRAALPENPLINVD 60  
Db 1 MAADAFGDRMEEPPLDRAPIYVAGFLALYDSGDELALDPDPTVRAALPDNPLINVD 60

Qy 61 HRARCEVGRVLAIVNDPRGPFVGLIACVQLERVLTAASAAIFERRGPALSREERLLYL 120  
Db 61 HRACEVGRVLAIVNDPRGPFVGLIACVQLERVLTAASAAIFERRGPALSREERLLYL 120

Qy 121 ITNYPVSLSLTKRGDEVPDRTLFAHVALCAIGRLGTIVYTQSLDAAIAFFRLDP 180  
Db 121 ITNYPVSLSLTKRGDEVPDRTLFAHVALCAIGRLGTIVYTQSLDAAIAFFRLDP 180

Qy 181 ATREGVRRRAEALALAGRTWAPGVGVEALTHLLSTAVNNMLDRWSLVAERROGIA 240  
Db 181 ASREGARRLAEEALALAGRTWAPGVGVEALTHLLSTAVNNMLDRWSLVAERROGIA 240

Qy 241 GHTYLAQSEKFKIWAESAEPAPRGYKTAGPAMDTSPAASVPAPO----VAVRAROVAS 296  
Db 241 GHTYLAQSEKFKIWAESAEPAPRGYKTAGPAMDTSPAASVPAPO----VAVRAROVAS 296

Qy 297 SSSSSFFPAPADNPNVSVASGAPAPPDGGSYLWIPASHYNQLVTQSGAPRHPPLTACGL 356  
Db 301 S-----PVLPMNPVPTSGTPAPAPDGGSYLWIPASHYNQLVTQSGAPRHPPLTACGL 354

Qy 357 P-AAGTVAVCHPGAGSPHYPPPPHYPGMLFAGSPLEAQIAALVGAIAADROAGGLP 415  
Db 355 PAAAGVAVGPHGAGLSQHYPPHVAHQYPGVLFSGSPLEAQIAALVGAIAADROAGGQP 414



Db 616 ASSQTVDASASTGLDFGRDDADIFVSQLMSAR 647

RESULT 5  
S52216  
viral proteinase - rabies virus  
C/Species: rabies virus  
C/Date: 07-May-1995 #sequence\_revision 21-Jul-1995 #text\_change 21-Aug-1998  
C/Accession: S52216  
R/Camacho, A.; Tabaro, E.  
submitted to the EMBL Data Library, June 1994  
A/Reference number: S52215  
A/Accession: S52216  
A/Status: preliminary  
A/Molecule type: genomic RNA  
A/Residues: 1-522 <CAM>  
A/Cross-references: EMBL:X79983  
A/Note: the source is designated as pseudorabies virus  
C/Superfamily: varicella-zoster virus gene 33 protein

Query Match 27.9%; Score 943; DB 2; Length 522;  
Best Local Similarity 39.5%; Pred. No. 5.3e-42;  
Matches 261; Conservative 61; Mismatches 155; Indels 184; Gaps 33;

Qy	20	PIYVAGFLAYDGDGPGELALDDPTVRAALPPENPLPINDVHRCRCVGRVLAVVNDPRG	79
Db	3	PVTVSGYALYDR-DGGELALTREIVAAALPPAGPLPINIDHRCARCIGAVLAVVDDRG	61
Qy	80	PFVGLTIACVQLERVLETAASAIFFRRGPALSRERLLYLITNPLSPVSLSTKR-RGDE	138
Db	62	PFELGVVNCPLQGAVLARAVAPDFGDMRP--SDEERLLYLLNPLSPSLSRRLPGD-	118
Qy	139	VPDRTLFAHVALCAIGRRLLGTIVTDSLDAIAIPFRHLDPATRCGRREAAEAEALALA	198
Db	119	-APDETIFAHVALCVIGRRVTIVVDASPERAVGPPELSAGRSELLARAAESP--A	174
Qy	199	GRTWAPGEVALTHTLSTAVNNMLRDRWSLVAERRRQAGIAGHTYLAQSEKPKIWGAES	258
Db	175	ERYVHMSDALTRALLSTAVNNMLLRDWELVAERRRRREAGVRAHTYLAQAT---MW---	227
Qy	259	APAPERYGKTGAPGAMDTSPAASVPAPQAVARQVASSSSSSFFPAPADMPVSAAGAP	318
Db	228	GLLPKSG---ASPGA-----RAQ-----CGHGSP	248
Qy	319	APPPPGDGYLWIPASHYNOLVTGQAPRHPPLTACGLPAAGTVAVGHGAGSPSPHYPPP	378
Db	249	AERTPGD-YIFVPAQYNQVNVQR-----PAP-----	275
Qy	379	PAHPYQMLPAGSPSLEAQIAALVGAIAADRQGGULPAAAGDHGIRGSAKRHRHEVQPE	438
Db	276	-----SLESQLGAIVSA-AMDRRRRSPSEP---RPPARKRY-----	310
Qy	439	YDCGRDEPDRDPFYYPGEARPEPRPVDSSRAARQASGPHTITIALVCAVTSLQOELAHMR	498
Db	311	:-:-:-DDYADQNAYYPGEAPP-----PRATSRV---VSSLRQREISHLR	346
Qy	499	A---RTHAP-YGYPYP-----VG-PYHHPHADTETPAQPPRY-PAKAVYLPPLP	540
Db	347	AQHVRYPTPYAPAPQLLPAGAVGHPPHPH-HAAGALYPPMYAPQPGHAPPSPVA	405
Qy	541	HIAP--PGPPLSGAVPP-----PSYP-----PVAVTPGPAPPL-----HQP-SPA	577
Db	406	HAVPALPGLPASRRCGFAHVPAQVVPQPPVVOAQFVAFAAAAPPPLRLQQRHAPAPV	465
Qy	578	HAHPP-PPPGPTPPPAASLPQPEAPCAEAGALVNASSAHVVDTARAADLTVSQMGS	636
Db	466	QAAAPRAPASAPQPPVQOASVSAPR-PTESPAPIDASSAA---VACQRGADIFVSQLMSQ	521
Qy	637	R	637
Db	522	R	522



Db 538 ASGGVP-----CAPGPGVGCEPQAPQVTVTHNGHQAAPQAGGGATGATAANVEQRQP 589  
 Qy 602 PGAEA 606  
 Db 590 EGGEA 594  
  
 RESULT 9  
 Q0BE3R  
 BVRf2 (EC-RF3) protein - human herpesvirus 4 (strain B95-8)  
 N:Contains: BVRf2 (EC-RF3a) protein  
 C:Species: human herpesvirus 4, Epstein-Barr virus  
 C:Date: 03-Aug-1984 #sequence revision 03-Aug-1984 #text\_change 09-Jul-2004  
 C:Accession: A03798; S33049; S33050; S33051  
 C:Bankier, A.T.; Deininger, P.L.; Farrell, P.J.; Barrell, B.G.  
 Mol. Biol. Med. 1, 21-45, 1983  
 A:Title: Sequence analysis of the 17,166 bp EcoRI fragment C of B95-8 Epstein-Barr virus  
 A:Reference number: A30465; MUID:85035713; PMID:6092825  
 A:Accession: A03798  
 A:Molecule type: DNA  
 A:Residues: 1-605 <BAN>  
 A:CROSS-references: UNIPROT:P03234; EMBL:V01555; NID:G59074; PIDN:CAA24801.1; PID:G13349  
 R:Baer, R.; Bankier, A.T.; Biggin, M.D.; Deininger, P.L.; Farrell, P.J.; Gibson, T.J.; H  
 Nature 310, 207-211, 1984  
 A:Title: DNA sequence and expression of the B95-8 Epstein-Barr virus genome.  
 A:References number: A03794; MUID:84270667; PMID:6087149  
 A:Contents: annotation; protein coding region  
 C:Superfamily: varicella-zoster virus gene 33 protein  
 F:261-605/Product: BVRf2 (EC-RF3a) protein #status predicted <BVR>  
  
 Query Match 11.0%; Score 372.5; DB 1; Length 605;  
 Best Local Similarity 24.8%; Pred. No. 2.2e-12;  
 Matches 166; Conservative 79; Mismatches 237; Indels 187; Gaps 28;  
  
 Qy 17 RAVPIYAGFLALYDSGPGE--LALDPTVRAALPENPLPINVDHRCCEVGRVLAVV 74  
 Db 3 QAPSVYCGFERPDA-PPKDACLDPLTVKSQLPKPLPLTVLVEHLPDAPVGSVFLY 61  
 Qy 75 NDRPGPFVGLTIACVQLERVLETA-ASAAIFERRGPAISREERLLYITNLPVSLSSTK 133  
 Db 62 QSRAGLFSAAITSGDFLSLLDSIYHDCIAQSORPLPREPKVEALHA-WLPSLSLAS- 119  
 Qy 134 RRGDEVPPD-----RTLFAHVALCAIGRLGTIVTYTSLDAAIAFRRHLDPATR 183  
 Db 120 ----LHPDIPTQTTADGGKLSFFDHVSI CALGRRRRTTAVYGTDLAWLKHFSDLPSIA 174  
 Qy 184 EGVREAAEALAGRTWAPGVEALTH--LLSTAVNNMLDRWSLVAERRQAGIAGH 242  
 Db 175 AQIENDANAAK-----RESCGPDHFLPLFKLIJAKADAGFLRNKRVETLRDQGVANI PAE 230  
 Qy 243 TYLOASEKFIWGABSAPEAPRGYKTPAGAMDTSAPASVPAPQAVARAVASSSSSSS 302  
 Db 231 SYLKASD-----APDLQKPDKALQSPPASTD-----PATMLSGNAGEGATACGSA 277  
 Qy 303 -----PPADAMPNVSAGAPAPPDPPGSGSYLWI 331  
 Db 278 AAGQDLISVRNTFTLLQTLNDLNKPPRTPLPYAAPLPFPFHQAIATAPSYGPGAGAVA 337  
 Qy 332 PASHYNQLVTQSGAPRHPLTACGLPAAGTVAGHPGA-----GPSPHYPPPPAHP 382  
 Db 338 PAGGY-----FTSPGGYYAGP-AGGDPGAFAMDHTVYHPHPH-PPPAYFG 381  
 Qy 383 YPGMLFAGPSPLEAQIAALVGAIAADROAGGLPAAAGDGHGIRSKRRRHEVEPEYDCG 442  
 Db 382 LFG-LFGPPPPVPPYGGSHL-----RADYVPAPS-----RSNKRKRDPEDEEGG 426  
 Qy 443 RDEPRDPFYPGEARPEPRPVDSRRRAQAQSGPHETITLVGAVTSLQELAHMART- 501  
 Db 427 -----LFEFE-----DATLYRKDIAGLSKSVNELQHTLQALRRRTL 462  
 Qy 502 ---HAPYGPYPVGP-YHH-----PHADTETPAQPPRYPAKAVYLPHPHAPPGPPLS 550  
 Db 463 SYGHTGVGVCQQGFCYTHSGSYGFQPHOSYEYPR-----YYPHP----- 502



```
Db 432 -----ERDASSD-EDMS-FPGEA-----DHGKARKELKAHGRDNNNSGSDAKGD 476
Qy 479 TITALGAVTSLOOE-LAHWRARTHAPVGPVPHHPTETPAQPPRYPAKAVYL 537
Db 477 RYDDIRALQELKREMLAVRQIAPRALLAQAQLATPVASPTTTTSHQAEASEPOASTAAA 536
Qy 538 PPHIAPP-----GPPLSGAVPP 555
Db 537 ASPSTASHGSKSAERGWVWNSCRVAPPLEAVNPP 571

RESULT 12
Q08B88
UL80 protein - human cytomegalovirus (strain AD169)
C;Contains: capsid assembly protein; viral proteinase (EC 3.4.21.-)
C;Species: human cytomegalovirus, human herpesvirus 5
A;Note: host Homo sapiens (man)
C;Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 09-Jul-2004
C;Accession: S09843; S51034; S51035
R;Chee, M.S.; Bankier, A.T.; Beck, S.; Bohnl, R.; Brown, C.M.; Cerny, R.; Horsnell, T.;
M.; Barrell, B.G.
Curr. Top. Microbiol. Immunol. 154, 125-169, 1990
A;Title: Analysis of the protein-coding content of the sequence of human cytomegalovirus
A;Reference number: S09749; MUID:90269039; PMID:2161319
A;Accession: S09843
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-708 <CHE>
A;Cross-references: UNIPROT:P16753; EMBL:X17403; NID:G59591; PIDN:CAA35353.1; PID:gl7808
A;Note: possible protein-coding frames are given
R;Stevens, J.T.; Mapelli, C.; Tsao, J.; Hail, M.; O'Boyle II, D.; Weinheimer, S.P.; Dila
Eur. J. Biochem. 226, 361-367, 1994
A;Title: In vitro proteolytic activity and active-site identification of the human cytom
A;Reference number: S51034; MUID:95094793; PMID:8001553
A;Accession: S51034
A;Molecule type: protein
A;Residues: 110-131,'X',133-134 <ST2>
K;Note: assembly protein release site (Ala-256-Ser-257) and maturation site (Ala-643-Ser
C;Superfamily: cytomegalovirus capsid assembly protein
C;Keywords: capsid assembly; hydrolyase; serine proteinase
F;1-256/Product: viral proteinase #status predicted <PMAY>
F;336-708/Product: capsid assembly #status predicted <CAP>
F;132/Active site: Ser #status experimental
F;256-257/Cleavage site: Ala-Ser (viral proteinase) #status experimental
F;643-644/Cleavage site: Ala-Ser (viral proteinase) #status experimental

Query Match 9.4%; Score 317.5; DB 1; Length 708;
Best Local Similarity 23.7%; Pred. No. 1.8e-09;
Matches 150; Conservative 67; Mismatches 196; Indels 219; Gaps 27;

Qy 20 PIYVAGFLALYD-SGDPGELALDPDVTVAALPEN-----PLINVDHRCARCEGRV 70
Db 13 PTVYVGFRLARYDQSPDEAEELLPRDVVEHLWHAQGGQSPSLVALPLNINHDDTAVVGHV 72
Qy 71 LAVVNDPRGPFVGLTACVQLERVLETAASAIFERRGPALS-REERLLVLTINYLPSVS 129
Db 73 AMQSVRDGLFCUGVTSRFLFIVRRASEKSELVGRVPSPQDPKKVEFLSGSVAGLS 132
Qy 130 LSKRRRGDEVP-----PDRTLFAHVALCAIGRRGLTIVTYDTSLDAAIAPFRHLDPAT 182
Db 133 LSS-RRCDVVEAATSLSGSETTFPKHVALCSVRRRGLTAVYGRDPENWTVQRPDLTAAD 191
Qy 183 REGVREAAEALAGRTWAPGEAL-----THLLSTAVNNMRLDRWSLVARRR 235
Db 192 RDGLR-----AQWRCGST---AVDASGDPFRSDSYGLLGNVDALYIRERLPLKYDKQ 243
Qy 236 QAQIA-GHTYLOASEKFKIWAESSAPAPERGYKTGAPGAMDTSPAASVPAPQVAVAROV 294
Db 244 LVGVTERESVKKAS-----VSPEACDIK--AASAERS 274
Qy 295 ASSSSSSSFPAPADMMNPFVSASGAPPPPGDGSYLMIPASHYNQLVTGOSAPRHPRLTAC 354
```

```
Db 275 GDSRSQAATPAAGARVPSSSPSPVEPP-----SPVQPP----- 308
Qy 355 GLPAAGCTVAVGHGAGSPSPHYPPPPA-----HPY-----PGMLFAGSPLEAQI 398
Db 309 ALPASPSVL---PAEGPPSLSPSEPAAASMSHPLSAVPAATAPPGATVAGASPVSLL 365
Qy 399 A-----ALVGAIAADROAGGLPAAAGDHGIRGSAKRRRHEVEQPEYDCGRDEPD 447
Db 366 AWPHDGVYLPKDAFFSILGASRSA--VPV----- 392
Qy 448 RDPYTPG-----EARPEPRPVDSRRRAQASGPHETITALTALVCAVTSLOQELAHMART 501
Db 393 ---MYPGAVAAPSPASAPLPLSPYPAS-----YCAPVVGVDQLA---ARH 432
Qy 502 HAPYGVPPVGPVGHHPHADTETPAQPPRYPAKAVYLPPIHAPPGLPSLGAIVPPSPVPV 561
Db 433 FADY-----VDPHY-----PGWGRRYE-----PAPSLHPSYP-----VPPP----- 463
Qy 562 AVTPGPAPPLHQPSAPAHAPPPPPPPGPTPPPA 593
Db 464 -----PSPAYYRRRDSFGGMDPEPS 483

RESULT 13
WZBEM4
gene 17 protein - saimiriine herpesvirus 1 (strain 11)
C;Species: saimiriine herpesvirus 1
A;Note: host Saimiri sciureus (common squirrel monkey)
C;Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 16-Jul-1999
C;Accession: H36807
R;Albrecht, J.
submitted to the EMBL Data Library, January 1992
A;Description: Primary structure of the herpesvirus saimiri genome.
A;Reference number: A36806
A;Accession: H36807
A;Molecule type: DNA
A;Residues: 1-475 <ALB>
A;Cross-references: GB:X64346; NID:G60320; PIDN:CAA45641.1; PID:G60339
R;Albrecht, J.C.; Nicholas, J.; Biller, D.; Cameron, K.R.; Biesinger, B.; Newman, C.; Wi
J. Virol. 66, 5047-5058, 1992
A;Title: Primary structure of the herpesvirus saimiri genome.
A;Reference number: A37309; MUID:92333688; PMID:1321287
A;Contents: annotation; possible protein-coding frames
A;Note: neither amino acid nor nucleotide sequence is given
C;Genetics:
A;Gene: 17
C;Superfamily: varicella-zoster virus gene 33 protein

Query Match 8.9%; Score 301.5; DB 1; Length 475;
Best Local Similarity 21.8%; Pred. No. 8.1e-09;
Matches 135; Conservative 80; Mismatches 230; Indels 173; Gaps 20;

Qy 21 IYVAGF--LALYDSGDPGELALDPDVTVAALPENPLINVDHRCARCEGRVLAIVNDPR 78
Db 4 VYVAGFVDVAVYKVDV-VLYNLDDVSKCLPTKPIPLNIEHLPESTIGHTGLYAVTH 62
Qy 79 GPFVGLIACVQLERVLETAASAIFERRGPALSREERLLYLTINYLPSVLSLTKRGDE 138
Db 63 GVFCVGVHSEKFLHLTENLFSNSCAQATSKFLPYQPLEMLHTLWLPALSLSLCPTAQ 122
Qy 139 VPPDRTLFAHVALCAIGRRGLTIVTYDTSLDAAIAPFRHLDPATREGVREAAEALALA 198
Db 123 NAANTNMFORVSLCALGRRRGTVAVYSMNLEDAISQFCSISQAEVENIVQDSKNVDINSL 182
Qy 199 GRTWAPGVBEALHTLLSTAVNNMRLDRWSLVARRRQAGIAGHTYLOASEKFKINGAES 258
Db 183 PK---PVFNIDPHILMAKIDAGFKDRLQLLTKDKGVAKIKKLTYLKASE----- 230
Qy 259 APAPERGYKTGAPGAMDTSPAASVPAPQVAVARQVASSSSSSFPAPADMMNPFVSASGAP 318
Db 231 -----IGKPVTEDIS-----EDMNQHI----- 248
```





GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 10, 2005, 18:57:09 ; Search time 199 Seconds  
(without alignments)  
1841.777 Million cell updates/sec

Title: US-10-623-429-9  
Perfect score: 3385  
Sequence: 1 MASAEMRERLEAPLPRAVP.....VNVDTARAADLFVSMQMSR 637

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt\_02.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3385	100.0	637	2 P89449	P89449 human herpe
2	3370.5	99.6	638	2 Q69527	Q69527 human herpe
3	3351.5	99.0	636	2 Q69104	Q69104 human herpe
4	2334	69.0	635	2 Q69087	Q69087 human herpe
5	2333	68.9	635	1 VP40_HHV11	P10210 human herpe
6	1820	53.8	329	2 P90341	P90341 human herpe
7	1729.5	51.1	608	2 Q806A3	Q806A3 cercopithec
8	1368.5	40.4	651	2 Q8JY97	Q8JY97 saimirine
9	1361	40.2	562	2 Q918P4	Q918P4 bovine herp
10	1131	33.4	247	2 O12271	O12271 human herpe
11	1079.5	31.9	646	1 VP40_EHVLB	P28936 equine herp
12	1079.5	31.9	646	2 Q6S6T7	Q6S6T7 equine herp
13	1079.5	31.9	646	2 AAS45919	Aas45919 equine he
14	1056	31.2	619	2 Q6X238	Q6X238 bovine herp
15	1056	31.2	619	2 AAR86135	Aar86135 bovine he
16	1056	31.2	647	2 Q39278	Q39278 equine herp
17	1010	29.8	621	1 VP40_BHV1C	P54817 bovine herp
18	1010	29.8	621	2 Q76PK0	Q76PK0 bovine herp
19	1010	29.8	621	2 CAB01600	Cab01600 bovine he
20	960	28.4	524	2 Q83417	Q83417 suid herpes
21	904	26.7	526	2 Q85027	Q85027 suid herpes
22	838.5	24.8	639	2 Q782R9	Q782R9 gallid herp
23	838.5	24.8	639	2 Q785G2	Q785G2 marek's dis
24	838.5	24.8	639	2 Q9YZ26	Q9YZ26 gallid herp
25	824.5	24.4	605	1 VP40_VZVD	P09286 varicella-z
26	824.5	24.4	605	2 Q6QCM2	Q6QCM2 human herpe
27	824.5	24.4	605	2 AAT07715	Aat07715 human her
28	824.5	24.4	605	2 AAT07791	Aat07791 human her
29	792	23.4	643	2 Q9DPR1	Q9DPR1 meleagrid h
30	787	23.2	663	2 Q9EGP2	Q9EGP2 meleagrid h
31	787	23.2	663	2 AAS01665	Aas01665 turkey he

RESULT 1

P89449	PRELIMINARY;	PRT;	637 AA.
ID	P89449		
AC	P89449;		
DT	01-MAY-1997 (Tremblrel. 03, Created)		
DT	01-MAY-1997 (Tremblrel. 03, Last sequence update)		
DT	01-OCT-2003 (Tremblrel. 25, Last annotation update)		
DE	Protease.		
GN	Name=UL26;		
OS	Human herpesvirus 2 (HHV-2) (Human herpes simplex virus 2).		
OC	Viruses; dsDNA viruses, no RNA stage; Herpesviridae;		
OC	Alphaherpesvirinae; Simplexvirus.		
OX	NCBI_TaxID=10310;		
RN	[1] _TaxID=10310;		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=HG52;		
RX	MEDLINE=87111457; PubMed=3027242;		
RA	McGeoch D.J., Moss H.W., McNab D., Frame M.C.;		
RT	"DNA sequence and genetic content of the HindIII 1 region in the short		
RT	unique component of the herpes simplex virus type 2 genome;		
RT	identification of the gene encoding glycoprotein G, and evolutionary		
RT	comparisons.;"		
RL	J. Gen. Virol. 68:19-38(1987).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=HG52;		
RX	MEDLINE=90278430; PubMed=2161906;		
RA	Everett R., Fenwick M.;		
RT	"Comparative DNA sequence analysis of the host shutoff genes of		
RT	different strains of herpes simplex virus: type 2 strain HG52 encodes		
RT	a truncated UL41 product.;"		
RL	J. Gen. Virol. 71:1387-1390(1990).		
RN	[3]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=HG52;		
RX	MEDLINE=92113549; PubMed=1662697;		
RA	McGeoch D.J., Cunningham C., McIntyre G., Dolan A.;		
RT	"Comparative sequence analysis of the long repeat regions and		
RT	adjoining parts of the long unique regions in the genomes of herpes		
RT	simplex viruses types 1 and 2.;"		
RL	J. Gen. Virol. 72:3057-3075(1991).		
RN	[4]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=HG52;		
RX	MEDLINE=92356101; PubMed=1322965;		
RA	Barnett B.C., Dolan A., Telford E.A.R., Davison A.J., McGeoch D.J.;		
RT	"A novel herpes simplex virus gene (UL49A) encodes a putative membrane		
RT	protein with counterparts in other herpesviruses.;"		
RL	J. Gen. Virol. 73:2167-2171(1992).		
RN	[5]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=HG52;		
RX	MEDLINE=92356101; PubMed=1322965;		
RA	Barnett B.C., Dolan A., Telford E.A.R., Davison A.J., McGeoch D.J.;		
RT	"A novel herpes simplex virus gene (UL49A) encodes a putative membrane		
RT	protein with counterparts in other herpesviruses.;"		
RL	J. Gen. Virol. 73:2167-2171(1992).		
RN	[5]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=HG52;		
RX	MEDLINE=92356101; PubMed=1322965;		
RA	Barnett B.C., Dolan A., Telford E.A.R., Davison A.J., McGeoch D.J.;		
RT	"A novel herpes simplex virus gene (UL49A) encodes a putative membrane		
RT	protein with counterparts in other herpesviruses.;"		
RL	J. Gen. Virol. 73:2167-2171(1992).		
RN	[5]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=HG52;		
RX	MEDLINE=92356101; PubMed=1322965;		
RA	Barnett B.C., Dolan A., Telford E.A.R., Davison A.J., McGeoch D.J.;		
RT	"A novel herpes simplex virus gene (UL49A) encodes a putative membrane		
RT	protein with counterparts in other herpesviruses.;"		
RL	J. Gen. Virol. 73:2167-2171(1992).		
RN	[5]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=HG52;		
RX	MEDLINE=92356101; PubMed=1322965;		
RA	Barnett B.C., Dolan A., Telford E.A.R., Davison A.J., McGeoch D.J.;		
RT	"A novel herpes simplex virus gene (UL49A) encodes a putative membrane		
RT	protein with counterparts in other herpesviruses.;"		
RL	J. Gen. Virol. 73:2167-2171(1992).		
RN	[5]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=HG52;		
RX	MEDLINE=92356101; PubMed=1322965;		
RA	Barnett B.C., Dolan A., Telford E.A.R., Davison A.J., McGeoch D.J.;		
RT	"A novel herpes simplex virus gene (UL49A) encodes a putative membrane		
RT	protein with counterparts in other herpesviruses.;"		
RL	J. Gen. Virol. 73:2167-2171(1992).		
RN	[5]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=HG52;		
RX	MEDLINE=92356101; PubMed=1322965;		
RA	Barnett B.C., Dolan A., Telford E.A.R., Davison A.J., McGeoch D.J.;		
RT	"A novel herpes simplex virus gene (UL49A) encodes a putative membrane		
RT	protein with counterparts in other herpesviruses.;"		
RL	J. Gen. Virol. 73:2167-2171(1992).		
RN	[5]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=HG52;		
RX	MEDLINE=92356101; PubMed=1322965;		
RA	Barnett B.C., Dolan A., Telford E.A.R., Davison A.J., McGeoch D.J.;		
RT	"A novel herpes simplex virus gene (UL49A) encodes a putative membrane		
RT	protein with counterparts in other herpesviruses.;"		
RL	J. Gen. Virol. 73:2167-2171(1992).		
RN	[5]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=HG52;		
RX	MEDLINE=92356101; PubMed=1322965;		
RA	Barnett B.C., Dolan A., Telford E.A.R., Davison A.J., McGeoch D.J.;		
RT	"A novel herpes simplex virus gene (UL49A) encodes a putative membrane		
RT	protein with counterparts in other herpesviruses.;"		
RL	J. Gen. Virol. 73:2167-2171(1992).		
RN	[5]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=HG52;		
RX	MEDLINE=92356101; PubMed=1322965;		
RA	Barnett B.C., Dolan A., Telford E.A.R., Davison A.J., McGeoch D.J.;		
RT	"A novel herpes simplex virus gene (UL49A) encodes a putative membrane		
RT	protein with counterparts in other herpesviruses.;"		
RL	J. Gen. Virol. 73:2167-2171(1992).		
RN	[5]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=HG52;		
RX	MEDLINE=92356101; PubMed=1322965;		
RA	Barnett B.C., Dolan A., Telford E.A.R., Davison A.J., McGeoch D.J.;		
RT	"A novel herpes simplex virus gene (UL49A) encodes a putative membrane		
RT	protein with counterparts in other herpesviruses.;"		
RL	J. Gen. Virol. 73:2167-2171(1992).		
RN	[5]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=HG52;		
RX	MEDLINE=92356101; PubMed=1322965;		
RA	Barnett B.C., Dolan A., Telford E.A.R., Davison A.J., McGeoch D.J.;		
RT	"A novel herpes simplex virus gene (UL49A) encodes a putative membrane		
RT	protein with counterparts in other herpesviruses.;"		
RL	J. Gen. Virol. 73:2167-2171(1992).		
RN	[5]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=HG52;		
RX	MEDLINE=92356101; PubMed=1322965;		
RA	Barnett B.C., Dolan A., Telford E.A.R., Davison A.J., McGeoch D.J.;		
RT	"A novel herpes simplex virus gene (UL49A) encodes a putative membrane		
RT	protein with counterparts in other herpesviruses.;"		
RL	J. Gen. Virol. 73:2167-2171(1992).		
RN	[5]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=HG52;		
RX	MEDLINE=92356101; PubMed=1322965;		
RA	Barnett B.C., Dolan A., Telford E.A.R., Davison A.J., McGeoch D.J.;		
RT	"A novel herpes simplex virus gene (UL49A) encodes a putative membrane		
RT	protein with counterparts in other herpesviruses.;"		
RL	J. Gen. Virol. 73:2167-2171(1992).		
RN	[5]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=HG52;		
RX	MEDLINE=92356101; PubMed=1322965;		
RA	Barnett B.C., Dolan A., Telford E.A.R., Davison A.J., McGeoch D.J.;		
RT	"A novel herpes simplex virus gene (UL49A) encodes a putative membrane		
RT	protein with counterparts in other herpesviruses.;"		
RL	J. Gen. Virol. 73:2167-2171(1992).		
RN	[5]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=HG52;		
RX	MEDLINE=92356101; PubMed=1322965;		
RA	Barnett B.C., Dolan A., Telford E.A.R., Davison A.J., McGeoch D.J.;		
RT	"A novel herpes simplex virus gene (UL49A) encodes a putative membrane		
RT	protein with counterparts in other herpesviruses.;"		
RL	J. Gen. Virol. 73:2167-2171(1992).		
RN	[5]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=HG52;		
RX	MEDLINE=92356101; PubMed=1322965;		
RA	Barnett B.C., Dolan A., Telford E.A.R., Davison A.J., McGeoch D.J.;		
RT	"A novel herpes simplex virus gene (UL49A) encodes a putative membrane		
RT	protein with counterparts in other herpesviruses.;"		
RL	J. Gen. Virol. 73:2167-2171(1992).		
RN	[5]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=HG52;		
RX	MEDLINE=92356101; PubMed=1322965;		
RA	Barnett B.C., Dolan A., Telford E.A.R., Davison A.J., McGeoch D.J.;		
RT	"A novel herpes simplex virus gene (UL49A) encodes a putative membrane		
RT	protein with counterparts in other herpesviruses.;"		
RL	J. Gen. Virol. 73:2167-2171(1992).		
RN	[5]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=HG52;		
RX	MEDLINE=92356101; PubMed=1322965;		
RA	Barnett B.C., Dolan A., Telford E.A.R., Davison A.J., McGeoch D.J.;		
RT	"A novel herpes simplex virus gene (UL49A) encodes a putative membrane		
RT	protein with counterparts in other herpesviruses.;"		
RL	J. Gen. Virol. 73:2167-2171(1992).		
RN	[5]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=HG52;		
RX	MEDLINE=92356101; PubMed=1322965;		
RA	Barnett B.C., Dolan A., Telford E.A.R., Davison A.J., McGeoch D.J.;		
RT	"A novel herpes simplex virus gene (UL49A) encodes a putative membrane		
RT	protein with counterparts in other herpesviruses.;"		
RL	J. Gen. Virol. 73:2167-2171(1992).		
RN	[5]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=HG52;		
RX	MEDLINE=92356101; PubMed=1322965;		
RA	Barnett B.C., Dolan A., Telford E.A.R., Davison A.J., McGeoch D.J.;		
RT	"A novel herpes simplex virus gene (UL49A) encodes a putative membrane		
RT	protein with counterparts in other herpesviruses.;"		
RL	J. Gen. Virol. 73:2167-2171(1992).		
RN	[5]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=HG52;		
RX	MEDLINE=92356101; PubMed=1322965;		
RA	Barnett B.C., Dolan A., Telford E.A.R., Davison A.J., McGeoch D.J.;		
RT	"A novel herpes simplex virus gene (UL49A) encodes a putative membrane		
RT	protein with counterparts in other herpesviruses.;"		
RL	J. Gen. Virol. 73:2167-2171(1992).		
RN	[5]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=HG52;		
RX	MEDLINE=92356101; PubMed=1322965;		
RA	Barnett B.C., Dolan A., Telford E.A.R., Davison A.J., McGeoch D.J.;		
RT	"A novel herpes simplex virus gene (UL49A) encodes a putative membrane		
RT	protein with counterparts in other herpesviruses.;"		
RL	J. Gen. Virol. 73:2167-2171(1992).		
RN	[5]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=HG52;		
RX	MEDLINE=92356101; PubMed=1322965;		
RA	Barnett B.C., Dolan A., Telford E.A.R., Davison A.J., McGeoch D.J.;		
RT	"A novel herpes simplex virus gene (UL49A) encodes a putative membrane		
RT	protein with counterparts in other herpesviruses.;"		
RL	J. Gen. Virol. 73:2167-2171(1992).		
RN	[5]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=HG52;		
RX	MEDLINE=92356101; PubMed=1322965;		
RA	Barnett B.C., Dolan A., Telford E.A.R., Davison A.J., McGeoch D.J.;		
RT	"A novel herpes simplex virus gene (UL49A) encodes a putative membrane		
RT	protein with counterparts in other herpesviruses.;"		
RL	J. Gen. Virol. 73:2167-2171(1992).		
RN	[5]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=HG52;		
RX	MEDLINE=92356101; PubMed=1322965;		
RA	Barnett B.C., Dolan A., Telford E.A.R., Davison A.J., McGeoch D.J.;		
RT	"A novel herpes simplex virus gene (UL49A) encodes a putative membrane		
RT	protein with counterparts in other herpesviruses.;"		
RL	J. Gen. Virol. 73:2167-2171(1992).		
RN	[5]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=HG52;		
RX	MEDLINE=92356101; PubMed=1322965;		
RA	Barnett B.C., Dolan A., Telford E.A.R., Davison A.J., McGeoch D.J.;		
RT	"A novel herpes simplex virus gene (UL49A) encodes a putative membrane		
RT	protein with counterparts in other herpesviruses.;"		
RL	J. Gen. Virol. 73:2167-2171(1992).		
RN	[5]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=HG52;		
RX	MEDLINE=92356101; PubMed=1322965;		
RA	Barnett B.C., Dolan A., Telford E.A.R., Davison A.J., McGeoch D.J.;		
RT	"A novel herpes simplex virus gene (UL49A) encodes a putative membrane		
RT	protein with counterparts in other herpesviruses.;"		
RL	J. Gen. Virol. 73:2167-2171(1992).		
RN	[5]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=HG52;		
RX	MEDLINE=92356101; PubMed=1322965;		
RA	Barnett B.C., Dolan A., Telford E.A.R., Davison A.J., McGeoch D.J.;		
RT	"A novel herpes simplex virus gene (UL49A) encodes a putative membrane		
RT	protein with counterparts in other herpesviruses.;"		
RL	J. Gen. Virol. 73:2167-2171(1992).		
RN	[5]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=HG52;		
RX	MEDLINE=92356101; PubMed=1322965;		
RA	Barnett B.C., Dolan A., Telford E.A.R., Davison A.J., McGeoch D.J.;		
RT	"A novel herpes simplex virus gene (UL49A) encodes a putative membrane		
RT	protein with counterparts in other herpesviruses.;"		
RL	J. Gen. Virol. 73:2167-2171(1992).		
RN	[5]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=HG52;		
RX	MEDLINE=92356101; PubMed=1322965;		
RA	Barnett B.C., Dolan A., Telford E.A.R., Davison A.J., McGeoch D.J.;		
RT	"A novel herpes simplex virus gene (UL49A) encodes a putative membrane		
RT	protein with counterparts in other herpesviruses.;"		
RL	J. Gen. Virol. 73:2167-2171(1992).		
RN	[5]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=HG52;		
RX	MEDLINE=92356101; PubMed=1322965;		
RA	Barnett B.C., Dolan A., Telford E.A.R., Davison A.J., McGeoch D.J.;		
RT	"A novel herpes simplex virus gene (UL49A) encodes a putative membrane		
RT	protein with counterparts in other herpesviruses.;"		
RL	J. Gen. Virol. 73:2167-2171(1992).		
RN	[5]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=HG52;		
RX	MEDLINE=92356101; PubMed=1322965;		
RA	Barnett B.C., Dolan A., Telford E.A.R., Davison A.J., McGeoch D.J.;		
RT	"A novel herpes simplex virus gene (UL49A) encodes a putative membrane		
RT	protein with counterparts in other herpesviruses.;"		
RL	J. Gen. Virol. 73:2167-2171(1992).		
RN	[5]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=HG52;		
RX	MEDLINE=92356101; PubMed=1322965;		
RA	Barnett B.C., Dolan A., Telford E.A.R., Davison A.J., McGeoch D.J.;		
RT	"A novel herpes simplex virus gene (UL49A) encodes a putative membrane		
RT	protein with counterparts in other herpesviruses.;"		
RL	J. Gen. Virol. 73:2167-2171(1992).		
RN	[5]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=HG52;		
RX	MEDLINE=92356101; PubMed=1322965;		
RA	Barnett B.C., Dolan A., Telford E.A.R., Davison A.J., McGeoch D.J.;		
RT	"A novel herpes simplex virus gene (UL49A) encodes a putative membrane		
RT	protein with counterparts in other herpesviruses.;"		
RL	J. Gen. Virol. 73:2167-2171(1992).		
RN	[5]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=HG52;		
RX	MEDLINE=92356101; PubMed=1322965;		
RA	Barnett B.C., Dolan A., Telford E.A.R., Davison A.J., McGeoch D.J.;		
RT	"A novel herpes simplex virus gene (UL49A) encodes a putative membrane		
RT	protein with counterparts in other herpesviruses.;"		
RL	J. Gen. Virol. 73:2167-2171(1992).		
RN	[5]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=HG52;		
RX	MEDLINE=92356101; PubMed=1322965;		
RA	Barnett B.C., Dolan A., Telford E.A.R., Davison A.J., McGeoch D.J.;		
RT	"A novel herpes simplex virus gene (UL49A) encodes a putative membrane		
RT	protein with counterparts in other herpesviruses.;"		
RL	J. Gen. Virol. 73:2167-2171(1992).		
RN	[5]		

```
DR EMBL; Z86099; CAB06750.1; -.
DR HSSP; Q69527; 1AT3.
DR MEROPS; S21.001; -.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004252; F:serine-type endopeptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001847; Peptidase S21.
DR Pfam; PF00716; Peptidase S21; I.
DR PRINTS; PR00236; Peptidase S21; I.
DR PROTEASE.
KW Protease.
SQ
SEQUENCE 637 AA; 66940 MW; 96B9D1F68CDAB9C CRC64;
Query Match 100.0%; Score 3385; DB 2; Length 637;
Best Local Similarity 100.0%; Pred. No. 4.2e-145;
Matches 637; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MASAEMRERLEAPLPDRAVPIYVAGFLALYDSGDPGELALDPDTVRAALPPENPLPINVD 60
Db 1 MASAEMRERLEAPLPDRAVPIYVAGFLALYDSGDPGELALDPDTVRAALPPENPLPINVD 60
QY 61 HRARCEVGRVLAVNDPRGPFVGLIACVQLSERVLETAASAAI FERRGPALSRERLLYL 120
Db 61 HRARCEVGRVLAVNDPRGPFVGLIACVQLSERVLETAASAAI FERRGPALSRERLLYL 120
QY 121 ITNYLPSVLSLTKRGDEVPDPRTLFAHVALCAIGRRIGTI VTYDTSLDAATAPRHLD 180
Db 121 ITNYLPSVLSLTKRGDEVPDPRTLFAHVALCAIGRRIGTI VTYDTSLDAATAPRHLD 180
QY 181 ATREGVRRERAAEALALAGRTWAPGVEALHTLLSTAVNNMLDRWSLVAERRRQAGIA 240
Db 181 ATREGVRRERAAEALALAGRTWAPGVEALHTLLSTAVNNMLDRWSLVAERRRQAGIA 240
QY 241 GHTYLQASEKFKI WGAESAPAPERGYKTGAPGAMDTSPAASVPAPQAVRARQVASSSSS 300
Db 241 GHTYLQASEKFKI WGAESAPAPERGYKTGAPGAMDTSPAASVPAPQAVRARQVASSSSS 300
QY 301 SSFPAPADMPNVSASGAPAPPPGSGSYLWIPASHYNQLVTGOSAPRHPPLTACGLPAA 360
Db 301 SSFPAPADMPNVSASGAPAPPPGSGSYLWIPASHYNQLVTGOSAPRHPPLTACGLPAA 360
QY 421 HGIRSAKRRRHEVEQPEYDCGRDEPDRFPYYPGEARPEPRVDSRRAARQASGPHETI 480
Db 421 HGIRSAKRRRHEVEQPEYDCGRDEPDRFPYYPGEARPEPRVDSRRAARQASGPHETI 480
QY 481 TALVGAVTSLQQLAHMRARTHAPYGPYPVGYHHPHADTETPAQPPRYPAKAVYLPPP 540
Db 481 TALVGAVTSLQQLAHMRARTHAPYGPYPVGYHHPHADTETPAQPPRYPAKAVYLPPP 540
QY 541 HTAPGPPPLSGAVPPPSYPVAVTPGPAPPLHQPSPAHAHPPPPPGTPPPPAASLPQPE 600
Db 541 HTAPGPPPLSGAVPPPSYPVAVTPGPAPPLHQPSPAHAHPPPPPGTPPPPAASLPQPE 600
QY 601 APGAEGALVNASSAAHVNDTARAADLFVSQMMGSR 637
Db 601 APGAEGALVNASSAAHVNDTARAADLFVSQMMGSR 637
RESULT 2
Q69527 ID Q69527 PRELIMINARY; PRT; 638 AA.
AC Q69527;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE UL26 protease.
OS Human herpesvirus 2 (HHV-2) (Human herpes simplex virus 2).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxID=10310;
```

```
[1]
SEQUENCE FROM N.A.
RA O'Shannessy D.J., O'Donnell K.C., Hellmig B., Shabon U., O'Brien S.,
RA Chamberlain P., Holmes S., Abdel-Meguid S.S., Debouck C.M.;
RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U49329; AAA92139.1; -.
DR PDB; 1AT3; X-ray; A/B=1-247.
DR MEROPS; S21.001; -.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004252; F:serine-type endopeptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001847; Peptidase S21.
DR Pfam; PF00716; Peptidase S21; I.
DR PRINTS; PR00236; HSCAPSIDP40.
KW Protease.
SQ
SEQUENCE 638 AA; 67028 MW; 992F230937BBAE89 CRC64;
Query Match 99.6%; Score 3370.5; DB 2; Length 638;
Best Local Similarity 99.7%; Pred. No. 1.9e-144;
Matches 636; Conservative 1; Mismatches 0; Indels 1; Gaps 1;
QY 1 MASAEMRERLEAPLPDRAVPIYVAGFLALYDSGDPGELALDPDTVRAALPPENPLPINVD 60
Db 1 MASAEMRERLEAPLPDRAVPIYVAGFLALYDSGDPGELALDPDTVRAALPPENPLPINVD 60
QY 61 HRARCEVGRVLAVNDPRGPFVGLIACVQLSERVLETAASAAI FERRGPALSRERLLYL 120
Db 61 HRARCEVGRVLAVNDPRGPFVGLIACVQLSERVLETAASAAI FERRGPALSRERLLYL 120
QY 121 ITNYLPSVLSLTKRGDEVPDPRTLFAHVALCAIGRRIGTI VTYDTSLDAATAPRHLD 180
Db 121 ITNYLPSVLSLTKRGDEVPDPRTLFAHVALCAIGRRIGTI VTYDTSLDAATAPRHLD 180
QY 181 ATREGVRRERAAEALALAGRTWAPGVEALHTLLSTAVNNMLDRWSLVAERRRQAGIA 240
Db 181 ATREGVRRERAAEALALAGRTWAPGVEALHTLLSTAVNNMLDRWSLVAERRRQAGIA 240
QY 241 GHTYLQASEKFKI WGAESAPAPERGYKTGAPGAMDTSPAASVPAPQAVRARQVASSSSS 299
Db 241 GHTYLQASEKFKI WGAESAPAPERGYKTGAPGAMDTSPAASVPAPQAVRARQVASSSSS 300
QY 300 SSFPAPADMPNVSASGAPAPPPGSGSYLWIPASHYNQLVTGOSAPRHPPLTACGLPAA 359
Db 300 SSFPAPADMPNVSASGAPAPPPGSGSYLWIPASHYNQLVTGOSAPRHPPLTACGLPAA 360
QY 360 GTVAYGHGAGSPHYPPPPPAHPYGMIFAGSPLEAQIAALVGAIAADRQAGGLPAAAG 419
Db 360 GTVAYGHGAGSPHYPPPPPAHPYGMIFAGSPLEAQIAALVGAIAADRQAGGLPAAAG 420
QY 420 DHGIRSAKRRRHEVEQPEYDCGRDEPDRFPYYPGEARPEPRVDSRRAARQASGPHET 479
Db 420 DHGIRSAKRRRHEVEQPEYDCGRDEPDRFPYYPGEARPEPRVDSRRAARQASGPHET 480
QY 480 ITALVGAVTSLQQLAHMRARTHAPYGPYPVGYHHPHADTETPAQPPRYPAKAVYLPPP 539
Db 480 ITALVGAVTSLQQLAHMRARTHAPYGPYPVGYHHPHADTETPAQPPRYPAKAVYLPPP 540
QY 540 PHIAPGPPPLSGAVPPPSYPVAVTPGPAPPLHQPSPAHAHPPPPPGTPPPPAASLPQPE 599
Db 540 PHIAPGPPPLSGAVPPPSYPVAVTPGPAPPLHQPSPAHAHPPPPPGTPPPPAASLPQPE 600
QY 600 EAPGAEGALVNASSAAHVNDTARAADLFVSQMMGSR 637
Db 600 EAPGAEGALVNASSAAHVNDTARAADLFVSQMMGSR 638
RESULT 3
Q69104 ID Q69104 PRELIMINARY; PRT; 636 AA.
AC Q69104;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
```

DE Protease (Fragment).  
 GN Name=ICP35;  
 OS Human herpesvirus 2 (HHV-2) (Human herpes simplex virus 2).  
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
 OC Alphaherpesvirinae; Simplexvirus.  
 OX NCBI\_TaxID=10310;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=G;  
 RA Steffy K.K.S., Schoen S., Chen C.-M.;  
 RT "Nucleotide sequence of the herpes simplex virus type 2 gene encoding  
 the protease and capsid protein ICP35.";  
 RL J. Gen. Virol. 0:0-0(1995).  
 DR EMBL; L37443; AAA67703.1; -.  
 DR MEROPS; S21.001; -.  
 DR GO; GO:0008233; F:Peptidase activity; IEA.  
 DR GO; GO:0004252; F:serine-type endopeptidase activity; IEA.  
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
 DR InterPro; IPR001847; Peptidase\_S21.  
 DR Pfam; PF00716; Peptidase\_S21; 1.  
 DR PRINTS; PR00236; HSCVAPSIDP40.  
 KW Protease.  
 KM NON\_TER  
 FT NON\_TER  
 SQ SEQUENCE 636 AA; 66785 MW; 817FF5814F7A48C3 CRC64;  
 Query Match 99.0%; Score 3351.5; DB 2; Length 636;  
 Best Local Similarity 99.4%; Pred. No. 1.3e-143;  
 Matches 633; Conservative 1; Mismatches 2; Indels 1; Gaps 1;  
 QY 1 MASAEMRELEAPLPDRAVPIYVAGFLALYDSGDELALDPDTVRAALPPENPLPINVD 60  
 DB 1 MASAEMRELEAPLPDRAVPIYVAGFLALYDSGDELALDPDTVRAALPPENPLPINVD 60  
 QY 61 HRACEVGRVLA VVNDPRGPFVGLIACVQLERVLETAASAAIFERRGPALSREERLLYL 120  
 DB 61 HRACEVGRVLA VVNDPRGPFVGLIACVQLERVLETAASAAIFERRGPALSREERLLYL 120  
 QY 121 ITNYPVSLSTKRRGDEVPDRTLFAHVALCAIGRLGTIVTYDTSLDAAIAPFRHLDP 180  
 DB 121 ITNYPVSLSTKRRGDEVPDRTLFAHVALCAIGRLGTIVTYDTSLDAAIAPFRHLDP 180  
 QY 181 ATREGVRRRAAEALAGRTWAPGVEALHTLLSTAVNNMLDRWSLVAERRRQAGIA 240  
 DB 181 ATREGVRRRAAEALAGRTWAPGVEALHTLLSTAVNNMLDRWSLVAERRRQAGIA 240  
 QY 241 GHTYLOASEKFKIWGAEAPAPRGYKGTGAPGAMDTSPAASVPAQVAVRARQVA 300  
 DB 241 GHTYLOASEKFKIWGAEAPAPRGYKGTGAPGAMDTSPAASVPAQVAVRARQVA 300  
 QY 301 SSFPAPADMPVSAAGAPAPPPGDSGLYWPASHYNQLVTGOSAPRHPPLTACGLPAAG 360  
 DB 301 SSFPAPADMPVSAAGAPAPPPGDSGLYWPASHYNQLVTGOSAPRHPPLTACGLPAAG 360  
 QY 361 TVAYGHGAGSPHYPPPPAHYPGMLFAGSPLEAQIAALVGAIAADRQAGGLPAAAGD 420  
 DB 361 TVAYGHGAGSPHYPPPPAHYPGMLFAGSPLEAQIAALVGAIAADRQAGGLPAAAGD 420  
 QY 421 HGIRGSAKRRRHEVEQPEYDCGRDEPDRTFPYYPGEARPEPRVDSRRARQASGPHEI 480  
 DB 421 HGIRGSAKRRRHEVEQPEYDCGRDEPDRTFPYYPGEARPEPRVDSRRARQASGPHEI 480  
 QY 481 TALVGAVTSLQQLAHMRARHAPYGPYPVGVPHHPHADTETPAOPPRYPKAVVLP 540  
 DB 481 TALVGAVTSLQQLAHMRARHAPYGPYPVGVPHHPHADTETPAOPPRYPKAVVLP 540  
 QY 541 HIAPPGPLSGAVPPSPYVAVTPGAPPLHQPSPAHAHPPPPPPGTPPPPAASLPQPE 600  
 DB 541 HIAPPGPLSGAVPPSPYVAVTPGAPPLHQPSPAHAHPPPPPPGTPPPPAASLPQPE 600  
 QY 601 APGAEGALVNASSAAHVNDTARAADLFVSQMMGSR 637  
 DB 601 APGAEGALVNASSAAHVNDTARAADLFVSQMMGSR 636

RESULT 4  
 Q69087 PRELIMINARY; PRT; 635 AA.  
 ID Q69087  
 AC Q69087;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE UL26.  
 OS Human herpesvirus 1 (HHV-1) (Human herpes simplex virus 1).  
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
 OC Alphaherpesvirinae; Simplexvirus.  
 OX NCBI\_TaxID=10298;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94308117; PubMed=8034621;  
 RA Darke P.L., Chen E., Hall D.L., Sardana M.K., Veloski C.A.,  
 RA Lafemina R.L., Shafer J.A., Kuo L.C.;  
 RT "Purification of active herpes simplex virus-1 protease expressed in  
 J. Biol. Chem. 269:18708-18711(1994).  
 RL EMBL; L32018; AAA5828.1; -.  
 DR PIR; A54862; A54862.  
 DR HSSP; Q69527; IAT3.  
 DR GO; GO:0008233; F:Peptidase activity; IEA.  
 DR GO; GO:0004252; F:serine-type endopeptidase activity; IEA.  
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
 DR InterPro; IPR001847; Peptidase\_S21.  
 DR Pfam; PF00716; Peptidase\_S21; 1.  
 DR PRINTS; PR00236; HSCVAPSIDP40.  
 KW Protease.  
 KM CHAIN 1 247 HSV-1 protease.  
 FT CHAIN 253 610 ICP35 assembly protein.  
 SQ SEQUENCE 635 AA; 66454 MW; 3C7768EE1496FDBF CRC64;  
 Query Match 69.0%; Score 2334; DB 2; Length 635;  
 Best Local Similarity 74.0%; Pred. No. 9.9e-98;  
 Matches 477; Conservative 30; Mismatches 120; Indels 18; Gaps 7;  
 QY 1 MASAEMRELEAPLPDRAVPIYVAGFLALYDSGDELALDPDTVRAALPPENPLPINVD 60  
 DB 1 MASAEMRELEAPLPDRAVPIYVAGFLALYDSGDELALDPDTVRAALPPENPLPINVD 60  
 QY 61 HRACEVGRVLA VVNDPRGPFVGLIACVQLERVLETAASAAIFERRGPALSREERLLYL 120  
 DB 61 HRACEVGRVLA VVNDPRGPFVGLIACVQLERVLETAASAAIFERRGPALSREERLLYL 120  
 QY 121 ITNYPVSLSTKRRGDEVPDRTLFAHVALCAIGRLGTIVTYDTSLDAAIAPFRHLDP 180  
 DB 121 ITNYPVSLSTKRRGDEVPDRTLFAHVALCAIGRLGTIVTYDTSLDAAIAPFRHLDP 180  
 QY 181 ATREGVRRRAAEALAGRTWAPGVEALHTLLSTAVNNMLDRWSLVAERRRQAGIA 240  
 DB 181 ASREGARRLAEEALAGRTWAPGVEALHTLLSTAVNNMLDRWSLVAERRRQAGIA 240  
 QY 241 GHTYLOASEKFKIWGAEAPAPRGYKGTGAPGAMDTSPAASVPAQVAVRARQVA 296  
 DB 241 GHTYLOASEKFKIWGAEAPAPRGYKGTGAPGAMDTSPAASVPAQVAVRARQVA 296  
 QY 297 SSSSSFPAPADMPVSAAGAPAPPPGDSGLYWPASHYNQLVTGOSAPRHPPLTACGL 356  
 DB 297 SSSSSFPAPADMPVSAAGAPAPPPGDSGLYWPASHYNQLVTGOSAPRHPPLTACGL 356  
 QY 301 P-----PVLPMNVPVTSSTPAPAPGDSGLYWPASHYNQLVAGHAAPQPHSAFGF 354  
 DB 301 P-----PVLPMNVPVTSSTPAPAPGDSGLYWPASHYNQLVAGHAAPQPHSAFGF 354  
 QY 357 P-AAGTVAYGHGAGSPHYPPPPAHYPGMLFAGSPLEAQIAALVGAIAADRQAGGLP 415  
 DB 355 PAAAGSVAYGHGAGLSQHPYHPVHAYQYGVLFSGSPLEAQIAALVGAIAADRQAGGP 414  
 QY 416 AAGDGHGIRGSAKRRRHEVEQPEYDCGRDEPDRTFPYYPGEARPEPRVDSRRARQAS 475  
 DB 415 -AAGDPGVNGSGKRRRYEAGPSYSCDQDEPDADYYPYYPGEARGA PRGVDSRRARHSPG 473  
 QY 476 PHETITALVGAVTSLQQLAHMRARHAPYGPYPVGVPHHPHADTE-TPAQPPRYPAKA 534



RC STRAIN=HG52;  
RA MEDLINE=87111457; PubMed=3027242;  
RX McGeoch D.J., Moss H.W., McNab D., Frame M.C.;  
RT "DNA sequence and genetic content of the HindIII 1 region in the short  
RT unique component of the herpes simplex virus type 2 genome;  
RT identification of the gene encoding glycoprotein G, and evolutionary  
RT comparisons.";  
RL J. Gen. Virol. 68:19-38(1987).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=HG52;  
RX MEDLINE=92113549; PubMed=1662697;  
RA McGeoch D.J., Cunningham C., McIntyre G., Dolan A.;  
RT "Comparative DNA sequence analysis of the host shutoff genes of  
RT different strains of herpes simplex virus: type 2 strain HG52 encodes  
RT a truncated UL41 product.";  
RL J. Gen. Virol. 71:1387-1390(1990).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=HG52;  
RX MEDLINE=92113549; PubMed=1662697;  
RA McGeoch D.J., Cunningham C., McIntyre G., Dolan A.;  
RT "Comparative sequence analysis of the long repeat regions and  
RT adjoining parts of the long unique regions in the genomes of herpes  
RT simplex viruses types 1 and 2.";  
RL J. Gen. Virol. 72:3057-3075(1991).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN=HG52;  
RX MEDLINE=92356101; PubMed=1322965;  
RA Barnett B.C., Dolan A., Telford E.A.R., Davison A.J., McGeoch D.J.;  
RT "A novel herpes simplex virus gene (UL49A) encodes a putative membrane  
RT protein with counterparts in other herpesviruses.";  
RL J. Gen. Virol. 73:2167-2171(1992).  
RN [5]  
RP SEQUENCE FROM N.A.  
RC STRAIN=HG52;  
RA Dolan A.;  
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; Z86099; CAB06751.1; -;  
SQ SEQUENCE 329 AA; 37836 MW; 59B0B428E63451E6 CRC64;

Query Match 53.8%; Score 1820; DB 2; Length 329;  
Best Local Similarity 100.0%; Pred. No. 8e-75;  
Matches 329; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 309 MNVPSASGAPAPPPGSGSYLWTPASHYNQVLTGQAPRHPPLTACGLPAAGTVAYGHG 368  
DB 1 MNVPSASGAPAPPPGSGSYLWTPASHYNQVLTGQAPRHPPLTACGLPAAGTVAYGHG 60  
QY 369 AGSPHYPPPPAHYPGMLFAGSPLEAQIAALVGAIAADROAGGLPAAAGDHGIRGSAK 428  
DB 61 AGSPHYPPPPAHYPGMLFAGSPLEAQIAALVGAIAADROAGGLPAAAGDHGIRGSAK 120  
QY 429 RRRHEVEQPEYDCGRDEPRDFFYPGCEARPEPRVDSRRAAQAQSGPHETITLVGAVT 488  
DB 121 RRRHEVEQPEYDCGRDEPRDFFYPGCEARPEPRVDSRRAAQAQSGPHETITLVGAVT 180  
QY 489 SLOQELAHMRARHAPYGPYPGYPHHPADTETPAQPRYPYAKAVYLPPIHAPGPP 548  
DB 181 SLOQELAHMRARHAPYGPYPGYPHHPADTETPAQPRYPYAKAVYLPPIHAPGPP 240  
QY 549 LSGAVPPPSYPVAVTPGAPPLHQPSPAHAHPPPPPGPTPPAASLPQEPAPGAEGA 608  
DB 241 LSGAVPPPSYPVAVTPGAPPLHQPSPAHAHPPPPPGPTPPAASLPQEPAPGAEGA 300  
QY 609 LVNASSAAHVNDTARAADLFVQMMGSR 637  
DB 301 LVNASSAAHVNDTARAADLFVQMMGSR 329

RESULT 7  
Q806A3

ID Q806A3 PRELIMINARY; PRT; 608 AA.  
AC Q806A3;  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
DE Capsid maturation protease.  
GN Names=UL26;  
OS Cercopithecine herpesvirus 1 (CeHV-1) (Simian herpes B virus).  
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
OC Alphaherpesvirinae; Simplexvirus.  
OX NCBI\_TaxID=10325;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22607624; PubMed=12721804;  
RA Ohesawa K., Black D.H., Sato H., Rogers K., Eberle R.;  
RT "Sequence and genetic arrangement of the UL region of the monkey B  
RT virus (Cercopithecine herpesvirus 1) genome and comparison with the UL  
RT region of other primate herpesviruses.";  
RL Arch. Virol. 148:989-997(2003).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=B2490;  
RX MEDLINE=22628476; PubMed=12743273;  
RA Porelygina L., Zhu L., Zurkunen H., Mills R., Borodovsky M.,  
RA Hilliard J.K.;  
RT "Complete sequence and comparative analysis of the genome of herpes B  
RT virus (Cercopithecine herpesvirus 1) from a rhesus monkey.";  
RL J. Virol. 77:6167-6177(2003).  
DR EMBL; AB096160; BAC58065.1; -;  
DR HSSP; P09286; 1VZV.  
DR GO; GO:0008233; F:peptidase activity; IEA.  
DR GO; GO:0004252; F:serine-type endopeptidase activity; IEA.  
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
DR InterPro; IPR001847; Peptidase\_S21.  
DR Pfam; PF00716; Peptidase\_S21; 1.  
DR PRINTS; PR00236; HSCAPSIDP40.  
KW Protease.  
SQ SEQUENCE 608 AA; 62683 MW; 16BA7A1C54FCAA3C CRC64;

Query Match 51.1%; Score 1729.5; DB 2; Length 608;  
Best Local Similarity 59.6%; Pred. No. 1.7e-70;  
Matches 401; Conservative 28; Mismatches 143; Indels 101; Gaps 24;

QY 1 MASAEMRERLEAPLPDRAVPIYVAGFLALYDSDGDELADPDVTRAAALPENPLINVD 60  
DB 1 MGPAADPEGPPGDDADRPVPIYVAGFLALYDSDGDELADPDVTRAAALPENPLINVD 60  
QY 61 HRARCEVGRVLAVVNDPRGPFVGLIACVQLERVLETAASAAIFERRGPALSREERLLYL 120  
DB 61 HRARCEVGRVLAVVNDPRGPFVGLIACVQLERVLETAASAAIFERRGPALSREERLLYL 120  
QY 121 ITNYLPSVSLTKRDEVPDPTLFAHVALCAIGRLGTIVTYDTSLOAAIAFPFRLDP 180  
DB 121 VTNYPVSLSTRLEEGETPDNTFLVHVALCAIGRLGTIVTYDTSLOAAIAFPFRLAP 180  
QY 181 ATREGVRRRAAEALAGRTWAPGVREALTHLTSTAVNNMLDRWSLVAERRRRQAGIA 240  
DB 181 DSREILRRRAAEALGLGGRVWTPGAEALTRALLSTAVNNMLDRWSLVAERRRRQAGIA 240  
QY 241 GHYTLQASEKFKIWGAESAAPAPERG-YKTGAPGAMDTSPAASVPAPQAVRARQVASSSS 299  
DB 241 GHYTLQASEKFKIWGAESAPAPERGPPDGVYKRSSGATPDGHASASAPR-----AHPD 291  
QY 300 SSSFPAPADMPV-SASGAPAPPPGDSGLYLMIPASHYNQVLTGQAPRHPPLTACGLPA 358  
DB 292 PSDVP-EAEMNPAPPAGAPTCKPPGDSGLYLMIPAAHYNLVAG-----HPPAPVFGAPA 345  
QY 359 AGTVAYGHGAGSPHYPPPPAHYPGMLFAGSPLEAQIAALVGAIAADROAGGLPAAA 418  
DB 346 P-PPAPGAPAYAPHHH-----AGVPGVFPFGSPLEAQIAALVGAIAADROAGGLP 396  
QY 419 GGHGIRGSAKRRRHEVEQPEYDCGRDEPRDFFYPGCEARPEPRPV-DS-----RRAARQA 473

```

Db      397  -----SGKRRYBAER---DCDYDDRD-DAPYYPGEARPPRVVPSGGRGRAPPSA 445
QY      474  -SGPHETITALYGAVTSLQOELAHMARHAPY-GY-PPVGPY-----PPVGPY----- 514
Db      446  GTGASETIALYGAVTSLQOELAHMARARA-APYGPYQAARFALGVAGEPPAPQVAQQWE 504
QY      515  -----HHPHADTETPAQPPRYPAKAVYLPPIPHIAPGPPPLSGAVPPSPYVAV-TP 565
Db      505  APAAGALAHPP-----PPPRH-----QPHATPANVPL-----VFGPQPCAVALLAQ 545
QY      566  GPAPPLHOPSAHAHPPPPPGPTPPPAASLPQPEAPGAEA-GALVNASSAAHNVNDTAR 624
Db      546  GPAPGAAPGGQCA-----PAAGAAPPAA-----EAGETEAGALVNASSAATHVDVDVGR 595
QY      625  AADLFVSQMGRS 637
Db      596  AADLFVSQMGRS 608

RESULT 8
Q8JY97 PRELIMINARY; PRT; 651 AA.
AC Q8JY97;
DT 01-OCT-2002 (TREMELrel. 22, Created)
DT 01-OCT-2002 (TREMELrel. 22, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Protease.
OS Saimirine herpesvirus 1 (strain MV-5-PSL) (SaHV-1) (Marmoset
  herpesvirus).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae;
OX NCBI_TaxID=10353;
RN [1]_
RP SEQUENCE FROM N.A.
RA Brehears M.A., Black D.H., Ritchey J.W., Eberle R.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY095366; AAM22798.1; -.
DR HSSP; Q69527; IAT3.
DR MEROPS; S21.001; -.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004252; F:serine-type endopeptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001847; Peptidase_S21.
DR Pfam; PF00716; Peptidase_S21; 1.
DR PRINTS; PRO0236; HSCVAPSIDP40.
KW Protease.
SQ SEQUENCE 651 AA; 68104 MW; CF62E03E2EFF3BB9 CRC64;

Query Match 40.4%; Score 1368.5; DB 2; Length 651;
Best Local Similarity 49.1%; Pred. No. 3.4e-54;
Matches 340; Conservative 54; Mismatches 182; Indels 117; Gaps 25;

QY      8  ERLEAPLPRAVP--LYVAGFLALYDSGDPGELALDPDTVRAALPENPLPINVDRARC 65
Db      11  EREDGAVDRTPTRMVYAGFLALYDSDDAEVLVADPVVAASLPAPPELPINIDHKTRC 70
QY      66  EVGRVLAVNDPRGPFVGLIACVQLERVLETAASAAIFERRGPALSREERLLYLITNYL 125
Db      71  SVGQVLAVNDARGPPFLGVACGQLEHLVLEAASGDI FGRGPPLTREERLLYLITNYL 130
QY      126  PSVSLTKRGDEVP---PDRTLFAHVALCAIGRLGTITVYDTSLDAAIAFRHLDPAT 182
Db      131  PSVSLSTRDDSGAPVVGNG--LFAHVALCAIGRLGTITVYDVTMEGAVAFRRGLAS 188
QY      183  REGVREAAEALAGRTWAPVEALHTLLSTAVNNMLDRWSLVAERRRQAGIAGH 242
Db      189  RGQVLDAVAEAVSGERTWAPGARALTTLTATAVNNMLDRWSVVAERRRQAGIAGH 248
QY      243  TYLQASEKFKINGASAPAPERYKTCGAPGAMDTs--PAASVPAPQVAVRQVASS--- 297
Db      249  TYLQASAKFGL--SDASP-----RADAESAADQRHIRADGTQAPESAAERREHPSLR 300

```

```

QY      298  -----SSSSFFAPADMNVPSASGAPA-PPPPGDGSLYLPASHYNQL 339
Db      301  PPSPPCDPSRIODGSLVPSAFPMVNP-SAGSGFLPKPPDGDGSLYLPASHYNNL 359
QY      340  VTGQ--SAPRHPLTACGLPAAGTVAYGHGAGSPSHYPYPPPAHPYPMGLFAGSPLEAQ 397
Db      360  VSSQAATAPHLPIYSAHQVAG--YGPWGVVPA--YVPVPA-----AGPSPLETQ 407
QY      398  IALVGAIAADROAGGL-----PAAAGDHGIRGSAKRHRHEVEQPEYD-----CGRDEPRD 449
Db      408  IALVGAIAADROAGGL-----PAAAGDHGIRGSAKRHRHEVEQPEYD-----CGRDEPRD 449
QY      450  FPYTPGE-----ARPEPREVDSR--RAAQAQSG-PH--ETITLVGAVTSLQOELAHMAR 500
Db      465  -PYTPGEVSGGGEGPGRDRRFGTATRSAAATPHSNETIALVGVASLQOELTHLSY 523
QY      501  THAPYGPYPVYPYHHPHADTETPAQPPRYPAKAVYLPPIPHIAPGPPPLSGAVPPSPY 560
Db      524  QHGAF-----VPSQAAAQVWTPRPYFAP-----AAAAQASHQL 556
QY      561  VAVTFGAPPLHOPSAHAHPPPPPGPTPPPAASL-----PQPEAPGA-----EAG 607
Db      557  QQAQFSCAPVTQTTPPAQVVPAAALFAAPPAAVQSLVGVGVGAPMEPRAGDAADVASLEAD 616
QY      608  A--LVNASSAAHNVNDTARAADLFVSQMGRS 637
Db      617  APPLHINASCTRTVTDNARASDAFVAQMMDR 649

RESULT 9
Q918P4 PRELIMINARY; PRT; 562 AA.
AC Q918P4;
DT 01-DEC-2001 (TREMELrel. 19, Created)
DT 01-DEC-2001 (TREMELrel. 19, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE UL26.
OS Bovine herpesvirus 2.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxID=10295;
RN [1]_
RP SEQUENCE FROM N.A.
RA May J.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF410478; AAL05039.1; -.
DR HSSP; Q69527; IAT3.
DR MEROPS; S21.001; -.
DR GO; GO:0004252; F:serine-type endopeptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001847; Peptidase_S21.
DR Pfam; PF00716; Peptidase_S21; 1.
DR PRINTS; PRO0236; HSCVAPSIDP40.
SQ SEQUENCE 562 AA; 59809 MW; 1E783D9743D90CBD CRC64;

Query Match 40.2%; Score 1361; DB 2; Length 562;
Best Local Similarity 51.1%; Pred. No. 6.5e-54;
Matches 327; Conservative 47; Mismatches 154; Indels 112; Gaps 21;

QY      17  RAVPIYVAGFLALYDSGDPGELALDPDTVRAALPENPLPINVDRHRCCEVGRVLAVND 76
Db      16  RAVPIYVAGFLALYEGGDELALPREVVSSALPSGFPVPIVNDHRAQCEVGRVLTIVDD 75
QY      77  PRGPFVGLIACVQLERVLETAASAAIFERRGPALSREERLLYLITNYLPSVSLTKR-- 134
Db      76  PRGPFVGLVSCAQLEEAASAAALFERRGPALSREERLLYLITNYLPSVSLSRRL 135
QY      135  RGDEVPDRTLFAHVALCAIGRLGTITVYDTSLDAAIAPPHLDPATREGVRRAAAE 194
Db      136  PGDTSWED--LFRHVALCVIGRLGTITVYDVSISGAVAPFOHLSSTSRDAAIRAEAA 193
QY      195  LALAGRTWAPVEALHTLLSTAVNNMLDRWSLVAERRRQAGIAGTLYQASEKFKIW 254

```

```

Db 194 LALRDWAPCGDALARTILLSTAVNSMLRRDRSLVSRERRRQAGIAGHTYLQASEAFGIW 253
Qy 255 GAESAPAPERGYKTGAPCAMDTSPAAPQAVAPQAVARQAVASSSSSSF-----PAP-AD 308
Db 254 GAGTCAFGSGYKEHA-----RSPRATA-----SPEAGSSGTFSSVQVIGAPLPTAH 300
Qy 309 MNFVSASGAPPP-----PPGDSYLWIPASHYNQVLTQSPAPRHPPLTACGLPAAAGTVAAG 365
Db 301 MN-----CGPFAAPQARPSGDDSYLWIPAAHYNQVAAQ---PNHCQQOQHMOFP---RVSHG 351
Qy 366 HPGAGPSPPHPPHPPV-----GMLFAGSPLEAQJALVGAJAADROAGGLPAAAGDH 421
Db 352 GPMGPPYGH-----PIYTPPYGLGAVAPGVSPLSTQIALVGAJAADROQATDRNAA--EL 405
Qy 422 GIRSAKRRRHEVBQBYDCGRDEP-----DRDFPYFGEARPEPRPVDSRRARQASGPH 477
Db 406 RSQGGKRRN-----DYDDDDGSPRHYGRDVPYFGEAPIRRAPEQRRA--VSPD 458
Qy 478 ETITALVGVNTSLOELAHMRARTHAPYGPYPVPGPYHHPHADTETPAQPPRYPAKAVYL 537
Db 459 DTITALIGAVSSLOELAHMRARSQVAVCAVP-----TTAPAPQP-----496
Qy 538 PPHIAPGPPGLSGAVPPSPVPPVAVTPGAPPLHQPSPAHAHPPPPPGPTPPPAASLP 597
Db 497 -----LPPSSQSAQTQG-----HQP-----QIPIQTTSATPQPVAAP 531
Qy 598 QPEAPGAEGALVNASSAAHVNDTARAADLFVSQMMGSR 637
Db 532 PPQ-----VDASGVAKVDVDRRAADLFVAHMMGSR 562

RESULT 10
O12271 PRELIMINARY; PRT; 247 AA.
M012271:
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Viral core protein.
OS Human herpesvirus 1 (HHV-1) (Human herpes simplex virus 1).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxID=10298;
RN [1]
SEQUENCE FROM N.A.
RP MEDLINE=97433272; PubMed=9288912;
RA Apeles H., Gottschalk U., Guntermann D., Hansen J., Maessen J.,
RA Schmidt E., Schneider K., Schneider M., Ruebsamen-Waigmann H.;
RT "Expression of natural and synthetic genes encoding herpes simplex
RT virus 1 protease in Escherichia coli and purification of the
RT protein."
RL Eur. J. Biochem. 247:890-895 (1997).
DR EMBL; Z97070; CAB09805.1; -.
DR HSP; Q69527; IAT3.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0004252; F:serine-type endopeptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001847; Peptidase S21.
DR Pfam; PF00716; Peptidase S21; I.
DR PRINTS; PR00236; HSCVAPSIDP40.
KW Protease.
SQ SEQUENCE 247 AA; 26619 MW; 0BED160C2DF0AD0C CRC64;

Query Match 33.4%; Score 1131; DB 2; Length 247;
Best Local Similarity 90.3%; Pred. No. 7e-44;
Matches 223; Conservative 8; Mismatches 16; Indels 0; Gaps 0;

Qy 1 MASAEMERLEAPLPDRAVPIYVAGFLALYDSGDPGELADPPTVRAALPENPLPINVD 60
Db 1 MAADAPQDRNEEPLDRAVPIYVAGFLALYDSGDPGELADPPTVRAALPENPLPINVD 60
Qy 61 HRARCEVGRVLAVNDPRGPFVGLIACVOLERVLETAASAAIFERRGPPLSREERLLYL 120

```

```

Db 61 HRARCEVGRVLAVNDPRGPFVGLIACVOLERVLETAASAAIFERRGPPLSREERLLYL 120
Qy 121 ITNLYPSVLSLTKRRGDEVPPDRTLFAHVALCAIGRLGTIVTYDTSLDAIAIPFRHLD 180
Db 121 ITNLYPSVLSLTKRGLGEAHPDRTLFAHVALCAIGRLGTIVTYDTGLDAAIAPFRHLS 180
Qy 181 ATREGVRRERAAEALAGRTWAPGVREALTHLLSTAVNNMMLRDRSLVAERRRQAGIA 240
Db 181 ASREGARRLAAEALAGSLGRTPWAPGVREALTHLLSTAVNNMMLRDRSLVAERRRQAGIA 240
Qy 241 GHTYLQA 247
Db 241 GHTYLQA 247

RESULT 11
VP40_EHV1B STANDARD; PRT; 646 AA.
ID VP40_EHV1B STANDARD; PRT; 646 AA.
AC P28936; Q69263;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Capsid protein P40 (Virus structural gene 35 protein) [Contains:
DE Capsid protein VP24 (Assemblin) (Protease) (EC 3.4.21.97); Capsid
DE protein VP22A].
GN Name=35;
OS Equine herpesvirus 1 (strain Ab4p) (EHV-1) (Equine abortion virus).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID=31520;
RN [1]
SEQUENCE FROM N.A.
RP MEDLINE=92295566; PubMed=1318606;
RA Telford E.A.R., Watson M.S., McBride K., Davison A.J.;
RT "The DNA sequence of equine herpesvirus-1."
RL Virology 189:304-316 (1992).
CC -I- FUNCTION: VP22A is a component of the capsid core involved in
CC processing and packaging of progeny DNA. VP24 is a protease which
CC can proteolytically cleave itself and VP22A at the C-terminus.
CC -I- CATALYTIC ACTIVITY: Cleaves -Ala-|-Ser- and -Ala-|-Ala- bonds in
CC the scaffold protein.
CC -I- SIMILARITY: Belongs to peptidase family S21.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; M86664; AAB02470.1; -.
DR EMBL; M86664; AAB02471.1; -.
DR PIR; I36798; WZBEC8.
DR HSP; Q69527; IAT3.
DR MEROPS; S21.001; -.
DR InterPro; IPR001847; Peptidase S21.
DR Pfam; PF00716; Peptidase S21; I.
DR PRINTS; PR00236; HSCVAPSIDP40.
KW Capsid assembly; Coat protein; Hydrolyase; Serine protease.
FT CHAIN 1 646
FT CHAIN 318 646
FT CHAIN 1 242
FT CHAIN 243 7622
FT PROPEP 7623 646
FT SITE 242 243
FT SITE 622 623
FT ACT_SITE 55 55
FT ACT_SITE 123 123
FT ACT_SITE 142 142
SQ SEQUENCE 646 AA; 68579 MW; 1C1C28C48116276 CRC64;

Query Match 31.9%; Score 1079.5; DB 1; Length 646;

```

		Best Local Similarity 40.3%; Pred. No. 3.6e-41;		Matches 278; Conservative 75; Mismatches 212; Indels 125; Gaps 20;	
QY	18	AVPIYVAGFLALYDSDPGELALDDPTVRAALPPENPLPINVDHRCARCEVGRVLAVNDP	77		
Db	12	SLPIYVAGYIALYDMGCGELTLTRETVAALPPASRLPINIDHRNGCVGSEVLIVDDA	71		
QY	78	RGPFVGLIACVQLERVLTAASAAIFERRGPALSRERLLYLTNYLPSVLSLTKRGD	137		
Db	72	RGPFGLIINCPQLGAVLATAGDPFFGELSGELSEQRLLYLVSNYLPASLSLRRRLGP	131		
QY	138	EYPPDRTLFAHVALCAIGRLGTTIVYDTSLSAAIAPFRHLDPATREGVRRRAAEALAL	197		
Db	132	DEEPDETLEFAHSLCVIGRRVGTIVYDATPENAVAPFKRLSPSSREELLITAREAQSL	191		
QY	198	A-GRTWAPGEALHTLLSTAVNNMLDRWSLVAERRRQAGIAGHTYLOASEKFKI---	253		
Db	192	GDAATWHLSEDTLTVLLSTAVNNMLLRNWLNRARRRREAGIEGHTYLOASASFGITNG	251		
QY	254	-----WGAESAPAPERGYKTCGAPGAMDTSPAASVPAQVAVRQVASS	298		
Db	252	CNKADFCGAELVDT--CGYKSG-----EKVHGAPYSRVTLGAKAFTSSSPNALPSSDN	302		
QY	299	-----SSSFPAPADMNVPVSAGAPAPPPPGDGSYLWIPASHYNQLVTGQSAPRHP	350		
Db	303	DKGGIGERTQKHISAMASSNPQTLNAGAPLVSGD--YILVPAQYNQLVVGQHT-SHPP	359		
QY	351	LTACGLPAAGTVAYGHGPGSPHYPPPAHPYPMGLFAGP-----SPLEAQIAALV	402		
Db	360	INAGPAPVTHAV-----PSQYIPPAVNSLMPMSYQAPPVWSVPHSANLEAQITALV	411		
QY	403	GAIAADRCAGGLPAAAGDHGIRGS-----AKRRHEVEQPEYDCGRDEPDR	448		
Db	412	GALAADRKA---TKGSDPHVIGSQSCSPPLSPQERRYARKRRHWDATTRD-----DL	462		
QY	449	DPFYPGEARPEPRVDSRAARQASGHETITLAVGAVTSLQQLAHMRA-----RT	501		
Db	463	EGIYYPGERSPRP---GERRAGR-----PSTTIADLMGAVSSLQEVSQLRAIQVTVAQPO	515		
QY	502	HAPYGPYPVGVYHHPHADTETPAOPPRYPYAKAVLPPPH-----IAPGPPLSGAVPPP	556		
Db	516	AAPAGLYKEI-----PAVPPQY-SQYIQIQOHAVSAIVAPQLPGI-----P	556		
QY	557	SYPPVAVTFGPAPPLHQPSFAHAHPPPPPGTTPPAASLPQPEAPGAEGALVNASSAA	616		
Db	557	SQFTQAVLAPQVPAGEAPGSAKVAAASTAPQAEQARAAPQOFEAVTSAAVLPTVTPQAS	616		
QY	617	HVNVDITARA-----ADLFVSQMMGSR 637			
Db	617	SQTVDSASASTGLEFGRDDADIFVSQMMGAR 646			
RESULT 12					
Q6S6T7					
ID	Q6S6T7	PRELIMINARY;		PRT;	646 AA.
AC	Q6S6T7;				
DT	05-JUL-2004 (TrEMBLrel. 27, Created)				
DT	05-JUL-2004 (TrEMBLrel. 27, Last sequence update)				
DT	01-OCT-2004 (TrEMBLrel. 28, Last annotation update)				
DE	Capsid protein (Protease).				
OS	Equine herpesvirus 1.				
OC	Viruses; dsDNA viruses, no RNA stage; Herpesviridae;				
OC	Alphaherpesvirinae; Varicellovirus.				
OX	NCBI_TaxID=10326;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=V592;				
RA	Davis-Poynter N.J., Nugent J., Birch-Machin I., Allen G.P.;				
RL	Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=Ab4;				
RX	MEDLINE=92295566; PubMed=1318606;				

		Query Match 31.9%; Score 1079.5; DB 2; Length 646;		Best Local Similarity 40.3%; Pred. No. 3.6e-41;			
		Matches 278; Conservative 75; Mismatches 212; Indels 125; Gaps 20;					
RA	Telford E.A., Watson M.S., McBride K., Davison A.J.;						
RT	"The DNA sequence of equine herpesvirus-1.,"						
RL	Virology 189:304-316 (1992).						
RN	[3]						
RP	SEQUENCE FROM N.A.						
RC	STRAIN=Ab4;						
RA	Davison A.J.;						
RL	Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.						
DR	EMBL; AY464052; AAS45919.1; --						
DR	EMBL; AY665713; AAT67292.1; --						
DR	InterPro; IPR001847; Peptidase S21.						
DR	Pfam; PF00716; Peptidase S21; 1.						
DR	PRINTS; PR00236; HSCCAPSIDP40.						
KW	Protease.						
SQ	SEQUENCE 646 AA; 68579 MW; 1C1C28C848116276 CRC64;						
QY	18	AVPIYVAGFLALYDSDPGELALDDPTVRAALPPENPLPINVDHRCARCEVGRVLAVNDP	77				
Db	12	SLPIYVAGYIALYDMGCGELTLTRETVAALPPASRLPINIDHRNGCVGSEVLIVDDA	71				
QY	78	RGPFVGLIACVQLERVLTAASAAIFERRGPALSRERLLYLTNYLPSVLSLTKRGD	137				
Db	72	RGPFGLIINCPQLGAVLATAGDPFFGELSGELSEQRLLYLVSNYLPASLSLRRRLGP	131				
QY	138	EYPPDRTLFAHVALCAIGRLGTTIVYDTSLSAAIAPFRHLDPATREGVRRRAAEALAL	197				
Db	132	DEEPDETLEFAHSLCVIGRRVGTIVYDATPENAVAPFKRLSPSSREELLITAREAQSL	191				
QY	198	A-GRTWAPGEALHTLLSTAVNNMLDRWSLVAERRRQAGIAGHTYLOASEKFKI---	253				
Db	192	GDAATWHLSEDTLTVLLSTAVNNMLLRNWLNRARRRREAGIEGHTYLOASASFGITNG	251				
QY	254	-----WGAESAPAPERGYKTCGAPGAMDTSPAASVPAQVAVRQVASS	298				
Db	252	CNKADFCGAELVDT--CGYKSG-----EKVHGAPYSRVTLGAKAFTSSSPNALPSSDN	302				
QY	299	-----SSSFPAPADMNVPVSAGAPAPPPPGDGSYLWIPASHYNQLVTGQSAPRHP	350				
Db	303	DKGGIGERTQKHISAMASSNPQTLNAGAPLVSGD--YILVPAQYNQLVVGQHT-SHPP	359				
QY	351	LTACGLPAAGTVAYGHGPGSPHYPPPAHPYPMGLFAGP-----SPLEAQIAALV	402				
Db	360	INAGPAPVTHAV-----PSQYIPPAVNSLMPMSYQAPPVWSVPHSANLEAQITALV	411				
QY	403	GAIAADRCAGGLPAAAGDHGIRGS-----AKRRHEVEQPEYDCGRDEPDR	448				
Db	412	GALAADRKA---TKGSDPHVIGSQSCSPPLSPQERRYARKRRHWDATTRD-----DL	462				
QY	449	DPFYPGEARPEPRVDSRAARQASGHETITLAVGAVTSLQQLAHMRA-----RT	501				
Db	463	EGIYYPGERSPRP---GERRAGR-----PSTTIADLMGAVSSLQEVSQLRAIQVTVAQPO	515				
QY	502	HAPYGPYPVGVYHHPHADTETPAOPPRYPYAKAVLPPPH-----IAPGPPLSGAVPPP	556				
Db	516	AAPAGLYKEI-----PAVPPQY-SQYIQIQOHAVSAIVAPQLPGI-----P	556				
QY	557	SYPPVAVTFGPAPPLHQPSFAHAHPPPPPGTTPPAASLPQPEAPGAEGALVNASSAA	616				
Db	557	SQFTQAVLAPQVPAGEAPGSAKVAAASTAPQAEQARAAPQOFEAVTSAAVLPTVTPQAS	616				
QY	617	HVNVDITARA-----ADLFVSQMMGSR 637					
Db	617	SQTVDSASASTGLEFGRDDADIFVSQMMGAR 646					

		Query Match 31.9%; Score 1079.5; DB 2; Length 646;		Best Local Similarity 40.3%; Pred. No. 3.6e-41;			
		Matches 278; Conservative 75; Mismatches 212; Indels 125; Gaps 20;					
RA	Telford E.A., Watson M.S., McBride K., Davison A.J.;						
RT	"The DNA sequence of equine herpesvirus-1.,"						
RL	Virology 189:304-316 (1992).						
RN	[3]						
RP	SEQUENCE FROM N.A.						
RC	STRAIN=Ab4;						
RA	Davison A.J.;						
RL	Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.						
DR	EMBL; AY464052; AAS45919.1; --						
DR	EMBL; AY665713; AAT67292.1; --						
DR	InterPro; IPR001847; Peptidase S21.						
DR	Pfam; PF00716; Peptidase S21; 1.						
DR	PRINTS; PR00236; HSCCAPSIDP40.						
KW	Protease.						
SQ	SEQUENCE 646 AA; 68579 MW; 1C1C28C848116276 CRC64;						
QY	18	AVPIYVAGFLALYDSDPGELALDDPTVRAALPPENPLPINVDHRCARCEVGRVLAVNDP	77				
Db	12	SLPIYVAGYIALYDMGCGELTLTRETVAALPPASRLPINIDHRNGCVGSEVLIVDDA	71				
QY	78	RGPFVGLIACVQLERVLTAASAAIFERRGPALSRERLLYLTNYLPSVLSLTKRGD	137				
Db	72	RGPFGLIINCPQLGAVLATAGDPFFGELSGELSEQRLLYLVSNYLPASLSLRRRLGP	131				
QY	138	EYPPDRTLFAHVALCAIGRLGTTIVYDTSLSAAIAPFRHLDPATREGVRRRAAEALAL	197				
Db	132	DEEPDETLEFAHSLCVIGRRVGTIVYDATPENAVAPFKRLSPSSREELLITAREAQSL	191				
QY	198	A-GRTWAPGEALHTLLSTAVNNMLDRWSLVAERRRQAGIAGHTYLOASEKFKI---	253				
Db	192	GDAATWHLSEDTLTVLLSTAVNNMLLRNWLNRARRRREAGIEGHTYLOASASFGITNG	251				
QY	254	-----WGAESAPAPERGYKTCGAPGAMDTSPAASVPAQVAVRQVASS	298				
Db	252	CNKADFCGAELVDT--CGYKSG-----EKVHGAPYSRVTLGAKAFTSSSPNALPSSDN	302				
QY	299	-----SSSFPAPADMNVPVSAGAPAPPPPGDGSYLWIPASHYNQLVTGQSAPRHP	350				
Db	303	DKGGIGERTQKHISAMASSNPQTLNAGAPLVSGD--YILVPAQYNQLVVGQHT-SHPP	359				
QY	351	LTACGLPAAGTVAYGHGPGSPHYPPPAHPYPMGLFAGP-----SPLEAQIAALV	402				
Db	360	INAGPAPVTHAV-----PSQYIPPAVNSLMPMSYQAPPVWSVPHSANLEAQITALV	411				
QY	403	GAIAADRCAGGLPAAAGDHGIRGS-----AKRRHEVEQPEYDCGRDEPDR	448				
Db	412	GALAADRKA---TKGSDPHVIGSQSCSPPLSPQERRYARKRRHWDATTRD-----DL	462				
QY	449	DPFYPGEARPEPRVDSRAARQASGHETITLAVGAVTSLQQLAHMRA-----RT	501				
Db	463	EGIYYPGERSPRP---GERRAGR-----PSTTIADLMGAVSSLQEVSQLRAIQVTVAQPO	515				
QY	502	HAPYGPYPVGVYHHPHADTETPAOPPRYPYAKAVLPPPH-----IAPGPPLSGAVPPP	556				
Db	516	AAPAGLYKEI-----PAVPPQY-SQYIQIQOHAVSAIVAPQLPGI-----P	556				
QY	557	SYPPVAVTFGPAPPLHQPSFAHAHPPPPPGTTPPAASLPQPEAPGAEGALVNASSAA	616				
Db	557	SQFTQAVLAPQVPAGEAPGSAKVAAASTAPQAEQARAAPQOFEAVTSAAVLPTVTPQAS	616				
QY	617	HVNVDITARA-----ADLFVSQMMGSR 637					
Db	617	SQTVDSASASTGLEFGRDDADIFVSQMMGAR 646					

		Query Match 31.9%; Score 1079.5; DB 2; Length 646;		Best Local Similarity 40.3%; Pred. No. 3.6e-41;			
		Matches 278; Conservative 75; Mismatches 212; Indels 125; Gaps 20;					
RA	Telford E.A., Watson M.S., McBride K., Davison A.J.;						
RT	"The DNA sequence of equine herpesvirus-1.,"						
RL	Virology 189:304-316 (1992).						
RN	[3]						
RP	SEQUENCE FROM N.A.						
RC	STRAIN=Ab4;						
RA	Davison A.J.;						
RL	Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.						
DR	EMBL; AY464052; AAS45919.1; --						
DR	EMBL; AY665713; AAT67292.1; --						
DR	InterPro; IPR001847; Peptidase S21.						
DR	Pfam; PF00716; Peptidase S21; 1.						
DR	PRINTS; PR00236; HSCCAPSIDP40.						
KW	Protease.						
SQ	SEQUENCE 646 AA; 68579 MW; 1C1C28C848116276 CRC64;						
QY	18	AVPIYVAGFLALYDSDPGELALDDPTVRAALPPENPLPINVDHRCARCEVGRVLAVNDP	77				
Db	12	SLPIYVAGYIALYDMGCGELTLTRETVAALPPASRLPINIDHRNGCVGSEVLIVDDA	71				
QY	78	RGPFVGLIACVQLERVLTAASAAIFERRGPALSRERLLYLTNYLPSVLSLTKRGD	137				
Db	72	RGPFGLIINCPQLGAVLATAGDPFFGELSGELSEQRLLYLVSNYLPASLSLRRRLGP	131				
QY	138	EYPPDRTLFAHVALCAIGRLGTTIVYDTSLSAAIAPFRHLDPATREGVRRRAAEALAL	197				
Db	132	DEEPDETLEFAHSLCVIGRRVGTIVYDATPENAVAPFKRLSPSSREELLITAREAQSL	191				
QY	198	A-GRTWAPGEALHTLLSTAVNNMLDRWSLVAERRRQAGIAGHTYLOASEKFKI---	253				
Db	192	GDAATWHLSEDTLTVLLSTAVNNMLLRNWLNRARRRREAGIEGHTYLOASASFGITNG	251				
QY	254	-----WGAESAPAPERGYKTCGAPGAMDTSPAASVPAQVAVRQVASS	298				
Db	252	CNKADFCGAELVDT--CGYKSG-----EKVHGAPYSRVTLGAKAFTSSSPNALPSSDN	302				
QY	299	-----SSSFPAPADMNVPVSAGAPAPPPPGDGSYLWIPASHYNQLVTGQSAPRHP	350				
Db	303	DKGGIGERTQKHISAMASSNPQTLNAGAPLVSGD--YILVPAQYNQLVVGQHT-SHPP	359				
QY	351	LTACGLPAAGTVAYGHGPGSPHYPPPAHPYPMGLFAGP-----SPLEAQIAALV	402				
Db	360	INAGPAPVTHAV-----PSQYIPPAVNSLMPMSYQAPPVWSVPHSANLEAQITALV	411				
QY	403	GAIAADRCAGGLPAAAGDHGIRGS-----AKRRHEVEQPEYDCGRDEPDR	448				
Db	412	GALAADRKA---TKGSDPHVIGSQSCSPPLSPQERRYARKRRHWDATTRD-----DL	462				
QY	449	DPFYPGEARPEPRVDSRAARQASGHETITLAVGAVTSLQQLAHMRA-----RT	501				
Db	463	EGIYYPGERSPRP---GERRAGR-----PSTTIADLMGAVSSLQEVSQLRAIQVTVAQPO	515				
QY	502	HAPYGPYPVGVYHHPHADTETPAOPPRYPYAKAVLPPPH-----IAPGPPLSGAVPPP	556				
Db	516	AAPAGLYKEI-----PAVPPQY-SQYIQIQOHAVSAIVAPQLPGI-----P	556				
QY	557	SYPPVAVTFGPAPPLHQPSFAHAHPPPPPGTTPPAASLPQPEAPGAEGALVNASSAA	616				
Db	557	SQFTQAVLAPQVPAGEAPGSAKVAAASTAPQAEQARAAPQOFEAVTSAAVLPTVTPQAS	616				
QY	617	HVNVDITARA-----ADLFVSQMMGSR 637					
Db	617	SQTVDSASASTGLEFGRDDADIFVSQMMGAR 646					

		Query Match 31.9%; Score 1079.5; DB 2; Length 646;		Best Local Similarity 40.3%; Pred. No. 3.6e-41;			
		Matches 278; Conservative 75; Mismatches 212; Indels 125; Gaps 20;					
RA	Telford E.A., Watson M.S., McBride K., Davison A.J.;						
RT	"The DNA sequence of equine herpesvirus-1.,"						
RL	Virology 189:304-316 (1992).						
RN	[3]						
RP	SEQUENCE FROM N.A.						
RC	STRAIN=Ab4;						
RA	Davison A.J.;						
RL	Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.						
DR	EMBL; AY464052; AAS45919.1; --						
DR	EMBL; AY665713; AAT67292.1; --						
DR	InterPro; IPR001847; Peptidase S21.						
DR	Pfam; PF00716; Peptidase S21; 1.						
DR	PRINTS; PR00236; HSCCAPSIDP40.						
KW	Protease.						
SQ	SEQUENCE 646 AA; 68579 MW; 1C1C28C848116276 CRC64;						
QY	18	AVPIYVAGFLALYDSDPGELALDDPTVRAALPPENPLPINVDHRCARCEVGRVLAVNDP	77				
Db	12	SLPIYVAGYIALYDMGCGELTLTRETVAALPPASRLPINIDHRNGCVGSEVLIVDDA	71				
QY	78	RGPFVGLIACVQLERVLTAASAAIFERRGPALSRERLLYLTNYLPSVLSLTKRGD	137				
Db	72	RGPFGLIINCPQLGAVLATAGDPFFGELSGELSEQRLLYLVSNYLPASLSLRRRLGP	131				
QY	138	EYPPDRTLFAHVALCAIGRLGTTIVYDTSLSAAIAPFRHLDPATREGVRRRAAEALAL	197				
Db	132	DEEPDETLEFAHSLCVIGRRVGTIVYDATPENAVAPFKRLSPSSREELLITAREAQSL	191				
QY	198	A-GRTWAPGEALHTLLSTAVNNMLDRWSLVAERRRQAGIAGHTYLOASEKFKI---	253				
Db	192	GDAATWHLSEDTLTVLLSTAVNNMLLRNWLNRARRRREAGIEGHTYLOASASFGITNG	251				
QY	254	-----WGAESAPAPERGYKTCGAPGAMDTSPAASVPAQVAVRQVASS	298				
Db	252	CNKADFCGAELVDT--CGYKSG-----EKVHGAPYSRVTLGAKAFTSSSPNALPSSDN	302				
QY	299	-----SSSFPAPADMNVPVSAGAPAPPPPGDGSYLWIPASHYNQLVTGQSAPRHP	350				
Db	303	DKGGIGERTQKHISAMASSNPQTLNAGAPLVSGD--YILVPAQYNQLVVGQHT-SHPP	359				
QY	351	LTACGLPAAGTVAYGHGPGSPHYPPPAHPYPMGLFAGP-----SPLEAQIAALV	402				
Db	360	INAGPAPVTHAV-----PSQYIPPAVNSLMPMSYQAPPVWSVPHSANLEAQITALV	411				
QY	403	GAIAADRCAGGLPAAAGDHGIRGS-----AKRRHEVEQPEYDCGRDEPDR	448				
Db	412	GALAADRKA---TKGSDPHVIGSQSCSPPLSPQERRYARKRRHWDATTRD-----DL	462				
QY	449	DPFYPGEARPEPRVDSRAARQASGHETITLAVGAVTSLQQLAHMRA-----RT	501				
Db	463	EGIYYPGERSPRP---GERRAGR-----PSTTIADLMGAVSSLQEVSQLRAIQVTVAQPO	515				
QY	502	HAPYGPYPVGVYHHPHADTETPAOPPRYPYAKAVLPPPH-----IAPGPPLSGAVPPP	556				
Db	516	AAPAGLYKEI-----PAVPPQY-SQYIQIQOHAVSAIVAPQLPGI-----P	556				
QY	557	SYPPVAVTFGPAPPLHQPSFAHAHPPPPPGTTPPAASLPQPEAPGAEGALVNASSAA	616				
Db	557	SQFTQAVLAPQVPAGEAPGSAKVAAASTAPQAEQARAAPQOFEAVTSAAVLPTVTPQAS	616				
QY	617	HVNVDITARA-----ADLFVSQMMGSR 637					
Db	617	SQTVDSASASTGLEFGRDDADIFVSQMMGAR 646					

		Query Match 31.9%; Score 1079.5; DB 2; Length 646;		Best Local Similarity 40.3%; Pred. No. 3.6e-41;			
		Matches 278; Conservative 75; Mismatches 212; Indels 125; Gaps 20;					
RA	Telford E.A., Watson M.S., McBride K., Davison A.J.;						
RT	"The DNA sequence of equine herpesvirus-1.,"						
RL	Virology 189:304-316 (1992).						
RN	[3]						
RP	SEQUENCE FROM N.A.						
RC	STRAIN=Ab4;						
RA	Davison A.J.;						
RL	Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.						
DR	EMBL; AY464052; AAS45919.1; --						
DR	EMBL; AY665713; AAT67292.1; --						
DR	InterPro; IPR001847; Peptidase S21.						
DR	Pfam; PF00716; Peptidase S21; 1.						
DR	PRINTS; PR00236; HSCCAPSIDP40.						
KW	Protease.						
SQ	SEQUENCE 646 AA; 68579 MW; 1C1C28C848116276 CRC64;						
QY	18	AVPIYVAGFLALYDSDPGELALDDPTVRAALPPENPLPINVDHRCARCEVGRVLAVND					



RESULT 14  
Q6X238  
ID Q6X238



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 10, 2005, 19:08:55 ; Search time 40 Seconds  
(without alignments)  
1056.114 Million cell updates/sec

Title: US-10-623-429-9  
Perfect score: 3385  
Sequence: 1 MASAEMRLEAPLPDRAVP.....VNVDTARAADLFVSMWGR 637

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA.\*  
1: /cgn2\_6/ptodata/1/iaa/5A COMB.pdp.\*  
2: /cgn2\_6/ptodata/1/iaa/5B COMB.pdp.\*  
3: /cgn2\_6/ptodata/1/iaa/6A COMB.pdp.\*  
4: /cgn2\_6/ptodata/1/iaa/6B COMB.pdp.\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS COMB.pdp.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pdp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2333	68.9	635	4	US-08-176-320-2
2	2323	68.6	635	1	US-07-832-855-2
3	1432	42.3	350	1	US-08-458-067-2
4	1432	42.3	350	5	PCT-US96-07795-2
5	1432	42.3	350	5	PCT-US96-07796-2
6	1262	37.3	306	1	US-08-279-754-2
7	1262	37.3	306	5	PCT-US95-09052-2
8	1252	37.0	247	3	US-09-167-434-4
9	1252	37.0	247	3	US-08-853-755-4
10	1126	33.3	247	3	US-09-167-434-3
11	1126	33.3	247	3	US-08-853-755-3
12	787	23.2	663	1	US-07-912-015-2
13	590.5	17.4	243	3	US-09-167-434-9
14	590.5	17.4	243	3	US-08-853-755-9
15	590.5	17.4	255	3	US-09-167-434-10
16	590.5	17.4	255	3	US-08-853-755-10
17	586.5	17.3	236	3	US-09-167-434-5
18	586.5	17.3	236	3	US-08-853-755-5
19	586.5	17.3	242	3	US-09-167-434-8
20	586.5	17.3	242	3	US-08-853-755-8
21	586.5	17.3	246	3	US-09-167-434-11
22	586.5	17.3	246	3	US-08-853-755-11
23	586.5	17.3	254	3	US-09-167-434-7
24	586.5	17.3	254	3	US-08-853-755-7
25	329.5	9.7	609	1	US-07-798-776-2
26	329.5	9.7	609	3	US-08-251-288A-2
27	329.5	9.7	609	3	US-09-298-819A-2

28	329.5	9.7	609	4	US-09-586-563C-2	Sequence 2, Appli
29	329.5	9.7	609	4	US-09-586-562C-2	Sequence 2, Appli
30	268.5	7.9	802	4	US-09-823-240A-2	Sequence 2, Appli
31	255.5	7.5	720	3	US-09-219-849-4	Sequence 4, Appli
32	255.5	7.5	777	1	US-08-642-255-53	Sequence 53, Appli
33	255	7.5	504	3	US-09-219-849-3	Sequence 3, Appli
34	255	7.5	561	1	US-08-642-255-52	Sequence 52, Appli
35	253.5	7.5	1184	4	US-09-266-225D-18	Sequence 18, Appli
36	253.5	7.5	1185	3	US-09-041-886-23	Sequence 23, Appli
37	253.5	7.5	1185	4	US-09-538-092-1209	Sequence 1209, Ap
38	248.5	7.3	1274	3	US-09-095-443-2	Sequence 2, Appli
39	242	7.1	408	1	US-07-609-716-65	Sequence 65, Appli
40	242	7.1	408	3	US-08-475-411A-65	Sequence 65, Appli
41	242	7.1	408	3	US-08-478-029A-65	Sequence 65, Appli
42	238.5	7.0	330	1	US-08-642-255-32	Sequence 32, Appli
43	236	7.0	960	3	US-09-219-849-5	Sequence 5, Appli
44	234.5	6.9	230	3	US-09-064-703-7	Sequence 7, Appli
45	229.5	6.8	424	4	US-09-538-092-1338	Sequence 1338, Ap

ALIGNMENTS

RESULT 1  
US-08-176-320-2  
; Sequence 2, Application US/08176320  
; Patent No. 6410704  
; GENERAL INFORMATION:  
; APPLICANT: Roizman, Bernard  
; APPLICANT: Liu, Fenyong  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE  
; TITLE OF INVENTION: PREPARATION AND USE OF A HERPES PROTEASE  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Alice O. Martin  
; STREET: 321 No. 6410704th Clark Street, Suite 800  
; CITY: Chicago  
; STATE: ILL  
; COUNTRY: USA  
; ZIP: 60610  
; COMPUTER READABLE FORM: disk  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/176,320  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/705,814  
; FILING DATE: 24-MAY-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Cooley, Ronald B  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (312)744-0090  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 635 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-08-176-320-2

Query Match 68.9%; Score 2333; DB 4; Length 635;  
Best Local Similarity 74.0%; Pred. No. 1.9e-160;  
Matches 477; Conservative 30; Mismatches 120; Indels 18; Gaps 7;  
Qy 1 MASAEMRLEAPLPDRAVPPIYVAGFLALYDSGDELALDPDTPVRAALPPENPLPND 60  
Db 1 MAADAFGDRMEELPDRAVPPIYVAGFLALYDSGDELALDPDTPVRAALPPENPLPND 60  
Qy 61 HRARCEVGRVLAVVNDPRGPFVGLIACVQLERVLTAASAAIFERRGPALSRERLLYL 120

Db 61 HRAGEVGRVAVDDPRGFFVGLIACVQLSERVLETAASAAIFERRGGPPLSREERLLYL 120  
QY 121 ITNYPVSLSTKRGDEVPDRITLFAHVALCAIGRRLLGTIVYDTSDDAAIAPPRHLD 180  
Db 121 ITNYPVSLSTKRGDEVPDRITLFAHVALCAIGRRLLGTIVYDTSDDAAIAPPRHLS 180  
QY 181 ATREGVRRRAAEALAGRTWAPGVLEALHTLLSTAVNNMRLDRWSLVAERROAGIA 240  
Db 181 ASREGARRLAEEALAGRTWAPGVLEALHTLLSTAVNNMRLDRWSLVAERROAGIA 240  
QY 241 GHTYLOASEKFKINGAESAPAPRGYKGTGAPCAMDTSPAASVPAQ----VAVRARQVAS 296  
Db 241 GHTYLOASEKFKINGAESAPAPRGYKGTGAPCAMDTSPAASVPAQ----VAVRARQVAS 300  
QY 297 SSSSSFFAPADMPNVPVSGAPAPPPGDSGLWIPASHYNQLVTGQSPRHPPLTACGL 356  
Db 301 S-----PVLPPMNPVPTSGTPAPAPPGDSGLWIPASHYNQLVAGHAAPQPHSAFGF 354  
QY 357 P-AAGTVAGHPCGAGSPHYPPPHAPHYPCMLFAGSPLEAQIAALVGNIAADROAGGLP 415  
Db 355 PAAAGSVAYGPHGAGLSQHQELAHMRARTSAPYGMVTPVAHYRQVGEPEPTTTHPALCPPEA 414  
QY 416 AAAGDHGIRGSAKRRRHEVEQPEYDCGRDEPRDPFYYPGEARPEPRPVDSSRAARQASG 475  
Db 415 -AAGDPGVKSGKRRRYEAGPSESYCDQDEPDADYPYYPGEARGAPRGVDSRRAARHSPG 473  
QY 476 PHETITLVGAVTSLQOELAHMRARTHAPYGPYPVGYHHPHADTE--TPAQPPRYPAKA 534  
Db 474 TNETITLMGAVTSLQOELAHMRARTSAPYGMVTPVAHYRQVGEPEPTTTHPALCPPEA 533  
QY 535 VYLPPLPHIAPPGLPLSGA--VPPPSYPVAVTPGAPPLHQSPAHAPPPPPPGPTPPP 592  
Db 534 VYRPPHSAFYGPPOGPASHAPTTPYAPAACTPGPPPP---PCPSTQTRAPLPTTEPAFP 590  
QY 593 AASLPQEPAPGAEGALVNASAAHVNVDTRAADLFVSQMMGSR 637  
Db 591 AATGQPEASNAEAGALVNASAAHVDVDTARAADLFVSQMMGAR 635

## RESULT 2

US-07-832-855-2  
; Sequence 2, Application US/07832855  
; Patent No. 547827  
; GENERAL INFORMATION:  
; APPLICANT: Roizman, Bernard  
; APPLICANT: Liu, Fenyong  
; TITLE OF INVENTION: Methods and Compositions of a  
; TITLE OF INVENTION: Preparation and Use of A Herpes Protease  
; NUMBER OF SEQUENCES: 15  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: ARNOLD, WHITE & DURKEE  
; STREET: 321 No. 5478727th Clark Street, Suite 800  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60610  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/832,855  
; FILING DATE: 19920207  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Cooley, Ronald B.  
; REGISTRATION NUMBER: 27,187  
; REFERENCE/DOCKET NUMBER: ARCD045  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (312) 744-0090  
; TELEFAX: (312) 245-4961

; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 635 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-07-832-855-2

Query Match 68.6%; Score 2323; DB 1; Length 635;  
Best Local Similarity 73.6%; Pred. No. 1e-159;  
Matches 475; Conservative 30; Mismatches 122; Indels 18; Gaps 7;

QY 1 MASAEEMRRLAPLPDRAVPIYVAGFLAYDSDGDELALDPDITVRAALPPENPLPIVD 60  
Db 1 MAADAPGRMEPLPDRAVPIYVAGFLAYDSDGDELALDPDITVRAALPPENPLPIVD 60  
QY 61 HRACEVGRVLA VNDPRGPFVGLIACVQLSERVLETAASAAIFERRGPALSREERLLYL 120  
Db 61 HRACEVGRVLA VNDPRGPFVGLIACVQLSERVLETAASAAIFERRGPALSREERLLYL 120  
QY 121 ITNYPVSLSTKRGDEVPDRITLFAHVALCAIGRRLLGTIVYDTSDDAAIAPPRHLD 180  
Db 121 ITNYPVSLSTKRGDEVPDRITLFAHVALCAIGRRLLGTIVYDTSDDAAIAPPRHLS 180  
QY 181 ATREGVRRRAAEALAGRTWAPGVLEALHTLLSTAVNNMRLDRWSLVAERROAGIA 240  
Db 181 ASREGARRLAEEALAGRTWAPGVLEALHTLLSTAVNNMRLDRWSLVAERROAGIA 240  
QY 241 GHTYLOASEKFKINGAESAPAPRGYKGTGAPCAMDTSPAASVPAQ----VAVRARQVAS 296  
Db 241 GHTYLOASEKFKINGAESAPAPRGYKGTGAPCAMDTSPAASVPAQ----VAVRARQVAS 300  
QY 297 SSSSSFFAPADMPNVPVSGAPAPPPGDSGLWIPASHYNQLVTGQSPRHPPLTACGL 356  
Db 301 S-----PVLPPMNPVPTSGTPAPAPPGDSGLWIPASHYNQLVAGHAAPQPHSAFGF 354  
QY 357 P-AAGTVAGHPCGAGSPHYPPPHAPHYPCMLFAGSPLEAQIAALVGNIAADROAGGLP 415  
Db 355 PAAAGSVAYGPHGAGLSQHQELAHMRARTHAPYGPYPVGYHHPHADTE--TPAQPPRYPAKA 414  
QY 416 AAAGDHGIRGSAKRRRHEVEQPEYDCGRDEPRDPFYYPGEARPEPRPVDSSRAARQASG 475  
Db 415 -AAGDPGVKSGKRRRYEAGPSESYCDQDEPDADYPYYPGEARGAPRGVDSRRAARHSPG 473  
QY 476 PHETITLVGAVTSLQOELAHMRARTHAPYGPYPVGYHHPHADTE--TPAQPPRYPAKA 534  
Db 474 TNETITLMGAVTSLQOELAHMRARTSAPYGMVTPVAHYRQVGEPEPTTTHPALCPPEA 533  
QY 535 VYLPPLPHIAPPGLPLSGA--VPPPSYPVAVTPGAPPLHQSPAHAPPPPPPGPTPPP 592  
Db 534 VYRPPHSAFYGPPOGPASHAPTTPYAPAACTPGPPPP---PCPSTQTRAPLPTTEPAFP 590  
QY 593 AASLPQEPAPGAEGALVNASAAHVNVDTRAADLFVSQMMGSR 637  
Db 591 PPTGSOPEASNAEAGALVNASAAHVDVDTARAADLFVSQMMGAR 635

## RESULT 3

US-08-458-067-2  
; Sequence 2, Application US/08458067  
; Patent No. 5728557  
; GENERAL INFORMATION:  
; APPLICANT: Register, Robert B.  
; APPLICANT: Shafer, Jules A.  
; TITLE OF INVENTION: HERPES SIMPLEX TYPE 1 PROTEASE MUTANTS  
; TITLE OF INVENTION: AND VECTORS  
; NUMBER OF SEQUENCES: 40  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Ms. Joanne M. Giesser  
; STREET: 126 East Lincoln Avenue, P.O. Box 2000-0907  
; CITY: Rahway  
; STATE: New Jersey  
; COUNTRY: US

```

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/07795
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Giesser, Joanne M.
; REGISTRATION NUMBER: 32,838
; REFERENCE/DOCKET NUMBER: 19457
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 594-3046
; TELEFAX: (908) 594-4720
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 350 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US96-07795-2

Query Match      42.3%; Score 1432; DB 5; Length 350;
Best Local Similarity 80.5%; Pred. No. 9,9e-96;
Matches 285; Conservative 13; Mismatches 46; Indels 10; Gaps 2

QY   1  MASAEWRERLEAPLPRAVPIYVAGFIALYDSGDPGELALDPTVTAAALPPENPLPINVD 60
    || :|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||
DB   1  MAADAPCDGRMEEP LPRAVPIYVAGFIALYDSGSGELALDPTVTAAALPPDNLPINVD 60

QY   61 HRARCEGVRLAVNDPRGPFFVGLIACVOLERVLETAASNAAIERRGPAISRERLLYL 120
    |||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||
DB   61 HRARCEGVRLAVVDPRGPFFVGLIACVOLERVLETAASNAAIERRGPPLSRERLLYL 120

QY   121 ITNYLPSVSLSXKRGDVEVPDRTLFAHVALCAIGRLGTITVTYDTSLDAAIAPFRHLDP 180
DB   121 ITNYLPSVSUATKRLGEAHPDRTLFAHVALCAIGRLGTITVTYDTGLDAAIAPFRHLSP 180

QY   181 ATREGVRRAAEAELALAGRTWAPGVREALTHLLSTAVNNMMLRDRLSVLAERRRQAGIA 240
    |||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||:|||
DB   181 ASREGARRLAAEAELALSGRTWAPGVREALTHLLSTAVNNMMLRDRLSVLAERRRQAGIA 240

QY   241 GHITYLOASEFKFWGAESAPAPERGYKTGAPGAMDTSPAASVPAQ----VAVRAQVAS 296
DB   241 GHITYLOASEFKFMWGAEFPVAPARGYKNGAPESTDIPPGSIATAAPQGDCRCIVRQGVAS 300

QY   297 SSSSSSFPADNMNYSASCAPAPPPPBGDSYLIWIPASHYNOLVTGOSAPRHPP 350
DB   301 P-----FVLPPMNVPVTSPTAPAPPBGDSYLIWIPASHYNQLVAGHAAPQPQ 348

RESULT 5
PCT-US96-07796-2
; Sequence 2, Application PC/TUS9607796
; GENERAL INFORMATION:
; APPLICANT: MERCK & CO., INC.
; APPLICANT: Registrar, Robert B.
; APPLICANT: Shafer, Jules A.
; TITLE OF INVENTION: HERPES SIMPLEX TYPE 1 PROTEASE MUTANTS
; TITLE OF INVENTION: AND VECTORS
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ms. Joanne M. Giesser
; STREET: 126 East Lincoln Avenue, P.O. Box 2000-0907
; CITY: Rahway
; STATE: New Jersey
; COUNTRY: US
; ZIP: 07065-0907
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30

```



```

; TELEPHONE: (908) 594-3046
; TELEFAX: (908) 594-4720
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 306 amino acids
; TYPE: amino acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
PCT-US95-09052-2

Query Match 37.3%; Score 1262; DB 5; Length 306;
Best Local Similarity 84.0%; Pred. No. 1.6e-83;
Matches 252; Conservative 11; Mismatches 33; Indels 4; Gaps 1;

QY 1 MASAENRERLEAPLPDRAPIYVAGFLALYDSDGPGELALDPTVRAALPPNPINVD 60
Db 1 MAADAFQDRMEELPDRAPIYVAGFLALYDSDGPGELALDPTVRAALPPNPINVD 60
QY 61 HRARCEVGRVLAIVNDPRGPFVGLIACVQLERVLETAASAAIFERRGPALESREERLLYL 120
Db 61 HRARCEVGRVLAIVNDPRGPFVGLIACVQLERVLETAASAAIFERRGPALESREERLLYL 120
QY 121 ITNLPVSLSTKRGDEVPDPRTLFAHVALCAIGRRLGITVYDTSLDAAIAPFRHLDP 180
Db 121 ITNLPVSLSTKRGDEVPDPRTLFAHVALCAIGRRLGITVYDTSLDAAIAPFRHLSP 180
QY 181 ATREGVRRREAAEALAGRTWAPGVLEALHTLLSTAVNNMLRDRWSLVAERRRQAGIA 240
Db 181 ASREGARRLAAEALAGRTWAPGVLEALHTLLSTAVNNMLRDRWSLVAERRRQAGIA 240
QY 241 GHTYLOASEKFKTWGAEAPAPRGYKGTAPGAMDTSPAASVPAPO----VAVRARQVAS 296
Db 241 GHTYLOASEKFKTWGAEAPAPRGYKGTAPGAMDTSPAASVPAPOQDCRCPIVRQGVAS 300

RESULT 8
US-09-167-434-4
; Sequence 4, Application US/09167434
; Patent No. 6008033
; GENERAL INFORMATION:
; APPLICANT: Abdel-Meguid, Sherin
; APPLICANT: Qiu, Xiayang
; APPLICANT: Culp, Jeffrey
; APPLICANT: Debouck, Christine
; APPLICANT: Janson, Cheryl A.
; APPLICANT: Hoog, Susan S.
; APPLICANT: Smith, Ward W.
; TITLE OF INVENTION: No. 6008033el Proteases, Compositions Capable of
; FILING DATE: Binding to Said Site, and Methods of Use Thereof
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESS: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road - P.O. Box 1539
; CITY: King of Prussia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19406-2799
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/167,434
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/853,755
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/022,470

```

```

; FILING DATE: 26-JUL-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/024,416
; FILING DATE: 21-AUG-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/030,901
; FILING DATE: 14-NOV-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/035,973
; FILING DATE: 21-JAN-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/039,191
; FILING DATE: 27-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Dinner, Dara L.
; REGISTRATION NUMBER: 33,680
; REFERENCE/DOCKET NUMBER: P50472-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5022
; TELEFAX: 610-270-5090
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 247 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-167-434-4

Query Match 37.0%; Score 1252; DB 3; Length 247;
Best Local Similarity 99.6%; Pred. No. 6.4e-83;
Matches 246; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MASAENRERLEAPLPDRAPIYVAGFLALYDSDGPGELALDPTVRAALPPNPINVD 60
Db 1 MASAENRERLEAPLPDRAPIYVAGFLALYDSDGPGELALDPTVRAALPPNPINVN 60
QY 61 HRARCEVGRVLAIVNDPRGPFVGLIACVQLERVLETAASAAIFERRGPALESREERLLYL 120
Db 61 HRARCEVGRVLAIVNDPRGPFVGLIACVQLERVLETAASAAIFERRGPALESREERLLYL 120
QY 121 ITNLPVSLSTKRGDEVPDPRTLFAHVALCAIGRRLGITVYDTSLDAAIAPFRHLDP 180
Db 121 ITNLPVSLSTKRGDEVPDPRTLFAHVALCAIGRRLGITVYDTSLDAAIAPFRHLDP 180
QY 181 ATREGVRRREAAEALAGRTWAPGVLEALHTLLSTAVNNMLRDRWSLVAERRRQAGIA 240
Db 181 ATREGVRRREAAEALAGRTWAPGVLEALHTLLSTAVNNMLRDRWSLVAERRRQAGIA 240
QY 241 GHTYLOA 247
Db 241 GHTYLOA 247

RESULT 9
US-08-853-755-4
; Sequence 4, Application US/08853755
; Patent No. 6083711
; GENERAL INFORMATION:
; APPLICANT: Abdel-Meguid, Sherin
; APPLICANT: Qiu, Xiayang
; APPLICANT: Culp, Jeffrey
; APPLICANT: Debouck, Christine
; APPLICANT: Janson, Cheryl A.
; APPLICANT: Hoog, Susan S.
; APPLICANT: Smith, Ward W.
; TITLE OF INVENTION: No. 6083711el Proteases, Compositions Capable of
; FILING DATE: Binding to Said Site, and Methods of Use Thereof
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESS: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road - P.O. Box 1539
; CITY: King of Prussia

```





Db 61 HRACCEVGRVLAVVDDPRGPFVGLIACVQLERVLETAASAAIFERRGPPLSREERLLYL 120  
 QY 121 ITNYPVSLSTKRGDEVPDRTLFAHVALCAIGRLGTVITVDTSLDAAIAPFRHLDP 180  
 Db 121 ITNYPVSLATKRLGGEAHPDRTLFAHVALCAIGRLGTVITVDTGLDAAIAPFRHLSP 180  
 QY 181 ATREGVRRRAAEALAGRTWAPGVEALHTLLSTAVNNMLRDRWSLVAERRRQAGIA 240  
 Db 181 ASREGARRLAEEALSGRTWAPGVEALHTLLSTAVNNMLRDRWSLVAERRRQAGIA 240  
 QY 241 GHTYLOA 247  
 Db 241 GHTYLOA 247

RESULT 11  
 US-08-853-755-3  
 ; Sequence 3, Application US/08853755  
 ; Patent No. 6083711  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Abdel-Meguid, Sherin  
 ; APPLICANT: Qiu, Xiayang  
 ; APPLICANT: Culp, Jeffrey  
 ; APPLICANT: Debouck, Christine  
 ; APPLICANT: Janson, Cheryl A.  
 ; APPLICANT: Hoog, Susan S.  
 ; APPLICANT: Smith, Ward W.  
 ; TITLE OF INVENTION: No. 6083711el Proteases, Compositions Capable of  
 ; TITLE OF INVENTION: Binding to Said Site, and Methods of Use Thereof  
 ; NUMBER OF SEQUENCES: 18  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: SmithKline Beecham Corporation  
 ; STREET: 709 Swedeland Road - P.O. Box 1539  
 ; CITY: King of Prussia  
 ; STATE: Pennsylvania  
 ; COUNTRY: USA  
 ; ZIP: 19406-2799  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/853,755  
 ; FILING DATE:  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 60/018,616  
 ; FILING DATE: 15-MAY-1996  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 60/022,470  
 ; FILING DATE: 26-JUL-1996  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 60/024,416  
 ; FILING DATE: 21-AUG-1996  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 60/030,901  
 ; FILING DATE: 14-NOV-1996  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 60/035,973  
 ; FILING DATE: 21-JAN-1997  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 60/039,191  
 ; FILING DATE: 27-FEB-1997  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Dinner, Dara L.  
 ; REGISTRATION NUMBER: 33,680  
 ; REFERENCE/DOCKET NUMBER: P50472-1  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 610-270-5022  
 ; TELEFAX: 610-270-5090  
 ; INFORMATION FOR SEQ ID NO: 3:  
 ; SEQUENCE CHARACTERISTICS:

; LENGTH: 247 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS:  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-08-853-755-3  
 Query Match 33.3%; Score 1126; DB 3; Length 247;  
 Best Local Similarity 89.9%; Pred. No. 8e-74;  
 Matches 222; Conservative 9; Mismatches 16; Indels 0; Gaps 0;  
 QY 1 MASAEMRERLEAPLPDRAVPYVAGFLALYDSGDPGELALDDPDTVRAALPPENPLPINVD 60  
 Db 1 MAADAPGDRMEELPDRAVPYVAGFLALYDSGDPGELALDDPDTVRAALPPENPLPINVN 60  
 QY 61 HRACCEVGRVLAVVNDPRGPFVGLIACVQLERVLETAASAAIFERRGPPLSREERLLYL 120  
 Db 61 HRACCEVGRVLAVVNDPRGPFVGLIACVQLERVLETAASAAIFERRGPPLSREERLLYL 120  
 QY 121 ITNYPVSLSTKRGDEVPDRTLFAHVALCAIGRLGTVITVDTSLDAAIAPFRHLDP 180  
 Db 121 ITNYPVSLATKRLGGEAHPDRTLFAHVALCAIGRLGTVITVDTGLDAAIAPFRHLSP 180  
 QY 181 ATREGVRRRAAEALAGRTWAPGVEALHTLLSTAVNNMLRDRWSLVAERRRQAGIA 240  
 Db 181 ASREGARRLAEEALSGRTWAPGVEALHTLLSTAVNNMLRDRWSLVAERRRQAGIA 240  
 QY 241 GHTYLOA 247  
 Db 241 GHTYLOA 247

RESULT 12  
 US-07-912-015-2  
 ; Sequence 2, Application US/07912015  
 ; Patent No. 5283191  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Morgan, Robin Wilson  
 ; APPLICANT: Willemse, Martha Jacoba  
 ; APPLICANT: Claessens, Johannes Antonius Joseph  
 ; APPLICANT: Sondermeijer, Paulus Jacobus Antonius  
 ; TITLE OF INVENTION: Marek's Disease Virus vaccine  
 ; NUMBER OF SEQUENCES: 4  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Organon Teknika Corporation  
 ; ADDRESSEE: Biotechnology Research Institute  
 ; STREET: 1330-A Piccard Drive  
 ; CITY: Rockville  
 ; STATE: Maryland  
 ; COUNTRY: U.S.A.  
 ; ZIP: 20850  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.24  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/07/912,015  
 ; FILING DATE:  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 07/615,211  
 ; FILING DATE: 16-NOV-1990  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Donna Bobrowicz  
 ; REGISTRATION NUMBER: 32,196  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (301)258-5200  
 ; INFORMATION FOR SEQ ID NO: 2:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 663 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear

```

; MOLECULE TYPE: protein
US-07-912-015-2

Query Match      23.2%; Score 787; DB 1; Length 663;
Best Local Similarity 31.6%; Pred. No. 7.5e-49;
Matches 229; Conservative 77; Mismatches 235; Indels 184; Gaps 26;

QY 21 IYVAGFLALYDSDGDELALDPDTVRAALPPENPLPINVDHRAECVGVYLAIVNDPRGP 80
Db :|||:||||: : : : :|||:||||:||||:||||:||||:||||:||||:||||:
9 VYVAGLYALYGADESDELNDKIDRAAIPTAPLPINIDHRRDCTCVGAVTALIDDEHGL 68
QY 81 FFVGLIACVOLERVLETAASAAIFERRGPALGREERLLYLIINYLPSVSLSKRGGDEVP 140
Db :|||:||||: : : : :|||:||||:||||:||||:||||:||||:||||:||||:
69 FFLGKINCVMVWRTLETAASQIFEL-DNLXDDKLLYIINYLPSVLSRRLAPGET 127
QY 141 PRTILFAHVALCAIGRLGTIVYDTSLSDAAIAPFRHLDPATREGVRRRAEALALAGR 200
Db :|||:||||: : : : :|||:||||:||||:||||:||||:||||:||||:||||:
128 ADETFLAHVALCLLGRIGTIVYDLTPEEATEPRKLPSPNSKATLLSQKETERLLGEM 187
QY 201 TWAPGVEALTHLLSTAVNNMLRDRWSLVARRRQAGIAGHTYLOASEKFKINGAESAP 260
Db :|||:||||: : : : :|||:||||:||||:||||:||||:||||:||||:||||:
188 VVYPSKNAITKALLGTAVNNMLLRWQIISERRRMAGITGQYLOASS----- 236
QY 261 APERGVKTCGAPGAMDTSPAASVPAP-----OVAVRAROVASS-----S 298
Db :|||:||||: : : : :|||:||||:||||:||||:||||:||||:||||:||||:
237 -----FTALTDSM-TSNNVSTHPIECENANPGNTQKDEMQVCISPAOTSETLNAGVL 289
QY 299 SSSSFPAPADMNPVSGAS-----GAPAPPPPGDGSYLMPASHYNNQLV----- 340
Db :|||:||||: : : : :|||:||||:||||:||||:||||:||||:||||:||||:
290 GCNDFHRLPHSDPASTSDQTNLSLIEPSMNTQSSRPFGD-DFIWVPIKSYNQLVSRNAS 348
QY 341 -----TQCSAPRHP-ITACGLPAAGTVAYGHGPGAGSPHYPPP-----PAH 381
Db :|||:||||: : : : :|||:||||:||||:||||:||||:||||:||||:||||:
349 QPTNIPDIAITSNQPPFIPPALMNTSISQHSIPSGYAOYG---YPTPVGTHNSLLPLG 404
QY 382 P-----YGMFLFAGSPLEAQIALYGAIAADROAGGLPAAAGD-HGIRGSA 427
Db :|||:||||: : : : :|||:||||:||||:||||:||||:||||:||||:||||:
405 PVNQMGFGYGFQVTP--LSYQGSPLKATLLECMTEKKRPVDEEHRGDDMHHTTREER 462
QY 428 KRRRHEVEQPEYDCGRDEP---DRDFF---YYPGEARP---BPRPVDSRRAARQASGPHE 478
Db :|||:||||: : : : :|||:||||:||||:||||:||||:||||:||||:||||:
463 GRR-----GRKRYFEDRSTESDLYYPGEFRSFPSPQASSMKYEETGGHRD 511
QY 479 -----TITALGAVTSLQELAHMART-----HAP-YGYPYPPVGPVH 515
Db :|||:||||: : : : :|||:||||:||||:||||:||||:||||:||||:||||:
512 LSQTGPVLSLGMGAVTSLQKEVERLNGNLPISNAQSSYGVNPGMHAPVYISYPPPGT-- 569
QY 516 HPHADTETPAQPRYPYPAKAVYLPPIHAPPGLPSGAVPPSPYVAVTPGPAPPLHQPS 575
Db :|||:||||: : : : :|||:||||:||||:||||:||||:||||:||||:||||:
570 HPTVSW-----PMGVERPMFSTEGKTSTNS---TVIPVPV----- 601
QY 576 PAHAHPPPPPPPGPTPPPAASLFPQEPAPGAEGAL-----VNASSAAHVNVDTPARAADLF 629
Db :|||:||||: : : : :|||:||||:||||:||||:||||:||||:||||:||||:
602 -----SDPEAGNVPITATISQERSDGIQKESIQSRDTMTNASAVAGIHRSTDAGVDVF 655
QY 630 VSQMM 634
Db :|||:||||: : : : :|||:||||:||||:||||:||||:||||:||||:||||:
656 INQMM 660

RESULT 13
US-09-167-434-9
; Sequence 9, Application US/09167434
; Patent No. 6008033
; GENERAL INFORMATION:
; APPLICANT: Abdel-Weguid, Sherin
; APPLICANT: Qiu, Xiyang
; APPLICANT: Culp, Jeffrey
; APPLICANT: DeBouck, Christine
; APPLICANT: Janson, Cheryl A.
; APPLICANT: Hoog, Susan S.
; APPLICANT: Smith, Ward W.
; TITLE OF INVENTION: No. 6008033el Proteases, Compositions Capable of
```

```

; TITLE OF INVENTION: Binding to Said Site, and Methods of Use Thereof
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road - P.O. Box 1539
; CITY: King of Prussia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19406-2799
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/167,434
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/853,755
; FILING DATE:
; APPLICATION NUMBER:
; APPLICATION NUMBER: US 60/022,470
; FILING DATE: 26-JUL-1996
; APPLICATION NUMBER: US 60/024,416
; FILING DATE: 21-AUG-1996
; APPLICATION DATA:
; APPLICATION NUMBER: US 60/030,901
; FILING DATE: 14-NOV-1996
; APPLICATION DATA:
; APPLICATION NUMBER: US 60/035,973
; FILING DATE: 21-JAN-1997
; APPLICATION DATA:
; APPLICATION NUMBER: US 60/039,191
; FILING DATE: 27-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Dinner, Dara L.
; REGISTRATION NUMBER: 33,680
; REFERENCE/DOCKET NUMBER: P50472-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5022
; TELEFAX: 610-270-5090
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 243 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-167-434-9

Query Match      17.4%; Score 590.5; DB 3; Length 243;
Best Local Similarity 50.0%; Pred. No. 3.5e-35;
Matches 124; Conservative 34; Mismatches 79; Indels 11; Gaps 3;

QY 1 MAAEVRERLEAPLPDRAVPIYVAGFLALYDSDGDELALDPDTVRAALPPENPLPINVD 60
Db :|||:||||: : : : :|||:||||:||||:||||:||||:||||:||||:||||:
1 MAAEADENCEA-----LYVAGLYALY-SKDEGELNITPEIVRSALPPTSKIPINID 51
QY 61 HRAECVGRVLAIVNDPRGPFFVGLIACVOLERVLETAASAAIFERRGPALSRERLLYL 120
Db :|||:||||: : : : :|||:||||:||||:||||:||||:||||:||||:||||:
52 HRKDCVGEVTAIEDIRGPFPLGVRCFQLHAVLFEAAHSNFFGNRDSVLSPLERAYL 111
QY 121 ITNYLPSVLSLTKRGGDEVPPDRTLFAHVALCAIGRLGTIVYDTSLSDAAIAPFRHLDP 180
Db :|||:||||: : : : :|||:||||:||||:||||:||||:||||:||||:||||:
112 VTNYLPSVLSLTKRGLSPNEIPDCNFFTHALCVGRRVGTVVNYDCTPESSEIEPRVLSM 171
QY 181 ATREGVRRRAEAEALALAGRTPWAPGVEALTHLLSTAVNNMLRDRWSLVARRRQAGIA 240
Db :|||:||||: : : : :|||:||||:||||:||||:||||:||||:||||:||||:
172 ESK--ARLLSLVKDYAGLNKVKVSEDKLAKVLLSTAVNNMLLRDRWDVAKRRREAGIM 229
QY 241 GHTYLOAS 248
```

```

QY 1 MASAMRERLEAPLPDRRAVPIYVAGFLYALYDSDGDELALDPDQTVRAALPPENPLPINVD 60
Db 1 MAEADSENCEBA-----LYVAGLYALY-SKDEGELNITPEIVRSALPPTSKIPINTD 51
QY 61 HRARCEVGRVLAVVNDPRGPFVGLIACVOLERVLETAASAAIPERRGPALSRERLLYL 120
Db 52 HRKDCVVGEVIAITEDIRGPFGLGVRCPQUHAVLFEAHSNFFGNRDSVLSPLERALLYL 111
QY 121 ITNYPVSVLSITKRRGDEVPPDRTLTFAHVALCAIGRRRLGTITVYDTSLSDAATAPFRHLPD 180
Db 112 VTNYLPSVLSLSSKRLSPNEIPDGNFFTHVALCVVGRRVGTVVNYDCTPESSIEPFRVLSM 171
QY 181 ATREGVRRBAEALALAGRTWAGVEALHTLTLSTAVNNMMLRDRWSLVABRRROAGIA 240
Db 172 ESK--ARLLSLVVDYAGLNVKWKVSEDKLAKVLLSTAVNNMLLRDRWDVAKRREAGIM 229
QY 241 GHTYLOAS 248
Db 230 GHVYLOAS 237

RESULT 15
US-09-167-434-10
; Sequence 10, Application US/09167434
; Patent No. 6008033
; GENERAL INFORMATION:
; APPLICANT: Abdel-Meguid, Sherin
; APPLICANT: Qiu, Xiayang
; APPLICANT: Culp, Jeffrey
; APPLICANT: Debouck, Christine
; APPLICANT: Janson, Cheryl A.
; APPLICANT: Hoog, Susan S.
; APPLICANT: Smith, Ward W.
; TITLE OF INVENTION: No. 6008033el Proteases, Compositions Capable of
; TITLE OF INVENTION: Binding to Said Site, and Methods of Use Thereof
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road - P.O. Box 1539
; CITY: King of Prussia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19406-2799
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/167.434
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/853,755
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/022,470
; FILING DATE: 26-JUL-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/024,416
; FILING DATE: 21-AUG-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/030,901
; FILING DATE: 14-NOV-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/035,973
; FILING DATE: 21-JAN-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/039,191
; FILING DATE: 27-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Dinner, Dara L.
; REGISTRATION NUMBER: 33,680

```

```

Db          230 GHVYLQAS 237

RESULT 14
US-08-853-755-9
; Sequence 9, Application US/08853755
; Patent No. 6083711
; GENERAL INFORMATION:
; APPLICANT: Abdel-Meguid, Sherin
; APPLICANT: Qiu, Xiayang
; APPLICANT: Culp, Jeffrey
; APPLICANT: Debouck, Christine
; APPLICANT: Janson, Cheryl A.
; APPLICANT: Hoog, Susan S.
; APPLICANT: Smith, Ward W.
; TITLE OF INVENTION: No. 6083711el Proteases, Compositions Capable of
; TITLE OF INVENTION: Binding to Said Site, and Methods of Use Thereof
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road - P.O. Box 1539
; CITY: King of Prussia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19406-2799
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/853,755
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/018,616
; FILING DATE: 15-MAY-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/022,470
; FILING DATE: 26-JUL-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/024,416
; FILING DATE: 21-AUG-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/030,901
; FILING DATE: 14-NOV-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/035,973
; FILING DATE: 21-JAN-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/039,191
; FILING DATE: 27-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Dinner, Dara L.
; REGISTRATION NUMBER: 33,680
; REFERENCE/DOCKET NUMBER: P50472-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5022
; TELEFAX: 610-270-5090
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 243 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-853-755-9

Query Match          17.4%; Score 590.5; DB 3; Length 243;
Best Local Similarity 50.0%; Pred. No. 3.5e-35;
Matches 124; Conservative 34; Mismatches 79; Indels 11; Gaps 3;

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 10, 2005, 19:18:39 ; Search time 148 Seconds  
(without alignments)

1551.918 Million cell updates/sec

Title: US-10-623-429-9

Perfect score: 3385

Sequence: 1 MASAEMRERLEAPLPDRAVE.....VNVDTRADLFVQMMSR 637

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1603904 seqs, 360571292 residues

Total number of hits satisfying chosen parameters: 1603904

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

1: /cgn2\_6/ptodata/1/pubpaa/PCT NEW PUB.pcp.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT NEW PUB.pcp.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW PUB.pcp.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pcp.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW PUB.pcp.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pcp.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW PUB.pcp.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pcp.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pcp.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pcp.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pcp.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW PUB.pcp.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pcp.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pcp.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pcp.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pcp.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US10E\_NEW PUB.pcp.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW PUB.pcp.\*  
19: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW PUB.pcp.\*  
20: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pcp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	3385	100.0	637	15	US-10-623-429-9
2	2333	68.9	635	14	US-10-214-932-86
3	351	10.4	536	17	US-10-779-597-30
4	294.5	8.7	534	16	US-10-437-963-159995
5	292	8.6	639	14	US-10-369-493-3962
6	284	8.4	825	16	US-10-408-765A-2285
7	281.5	8.3	483	16	US-10-437-963-191498
8	279.5	8.3	569	16	US-10-437-963-149276
9	275.5	8.1	426	16	US-10-437-963-142229
10	273	8.1	506	16	US-10-437-963-126337
11	271.5	8.0	521	16	US-10-437-963-179715
12	270	8.0	605	16	US-10-437-963-195106
13	268.5	7.9	373	16	US-10-437-963-149274

14	268.5	7.9	802	9	US-09-823-240-2	Sequence 2, Appli
15	267	7.9	280	17	US-10-425-115-231956	Sequence 231956,
16	267	7.9	445	16	US-10-437-963-161789	Sequence 161789,
17	266.5	7.9	469	16	US-10-437-963-136216	Sequence 136216,
18	266.5	7.9	497	16	US-10-437-963-122257	Sequence 122257,
19	266	7.9	501	16	US-10-437-963-161137	Sequence 161137,
20	265	7.8	289	16	US-10-437-963-172302	Sequence 172302,
21	263.5	7.8	563	16	US-10-437-963-198755	Sequence 198755,
22	263	7.8	408	16	US-10-437-963-184023	Sequence 184023,
23	263	7.8	554	16	US-10-437-963-174151	Sequence 174151,
24	262.5	7.8	639	14	US-10-369-493-2275	Sequence 2275, Ap
25	262.5	7.8	1405	16	US-10-437-963-156135	Sequence 156135,
26	261.5	7.7	507	16	US-10-437-963-124879	Sequence 124879,
27	261	7.7	285	16	US-10-437-963-189482	Sequence 189482,
28	261	7.7	1567	9	US-09-835-232-2	Sequence 2, Appli
29	261	7.7	1567	14	US-10-308-485-2	Sequence 2, Appli
30	260.5	7.7	554	16	US-10-437-963-119377	Sequence 119377,
31	258.5	7.6	442	16	US-10-437-963-197867	Sequence 197867,
32	258	7.6	431	16	US-10-437-963-204963	Sequence 204963,
33	256	7.6	500	16	US-10-437-963-139732	Sequence 139732,
34	255.5	7.5	720	14	US-10-342-331-4	Sequence 4, Appli
35	255.5	7.5	1023	10	US-09-893-519A-14	Sequence 14, Appli
36	255	7.5	504	14	US-10-342-331-3	Sequence 3, Appli
37	254.5	7.5	514	16	US-10-437-963-111701	Sequence 111701,
38	254.5	7.5	574	14	US-10-168-097A-76	Sequence 76, Appl
39	254.5	7.5	574	14	US-10-239-431A-38	Sequence 38, Appl
40	254	7.5	533	16	US-10-437-963-167716	Sequence 167716,
41	253.5	7.5	3122	14	US-10-200-562-201	Sequence 201, App
42	253.5	7.5	3122	14	US-10-237-551-201	Sequence 201, App
43	253.5	7.5	3122	14	US-10-237-551-250	Sequence 250, App
44	252.5	7.5	376	16	US-10-437-963-154970	Sequence 154970,
45	252	7.4	903	14	US-10-156-761-11093	Sequence 11093, A

#### ALIGNMENTS

#### RESULT 1

US-10-623-429-9  
; Sequence 9, Application US/10623429  
; Publication No. US20040072152A1  
; GENERAL INFORMATION:  
; APPLICANT: DAVID M. KOELLE  
; APPLICANT: ZHI LIU  
; APPLICANT: LAWRENCE COREY  
; TITLE OF INVENTION: RAPID, EFFICIENT PURIFICATION OF  
; TITLE OF INVENTION: HSV-SPECIFIC T-LYMPHOCYTES AND HSV ANTIGENS IDENTIFIED VIA  
; FILE OF INVENTION: SAME  
; FILE REFERENCE: 30967.11USUI  
; CURRENT APPLICATION NUMBER: US/10/623,429  
; CURRENT FILING DATE: 2003-07-18  
; PRIOR APPLICATION NUMBER: 60/396,791  
; PRIOR FILING DATE: 2002-07-18  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 9  
; LENGTH: 637  
; TYPE: PRT  
; ORGANISM: Herpes Simplex Virus 2  
US-10-623-429-9

Query Match 100.0%; Score 3385; DB 15; Length 637;  
Best Local Similarity 100.0%; Pred. No. 1.1e-199;  
Matches 637; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MASAEMRERLEAPLPDRAVEIYVAGFLALYDSGDPGELALDPTVRAALPPENPLPINV	60
Db	1	MASAEMRERLEAPLPDRAVEIYVAGFLALYDSGDPGELALDPTVRAALPPENPLPINV	60
Qy	61	HRARCEVGRVLA VVNDPRGPFVGLIACVOLERVLETAASAAIFERRGPALSRERLLYL	120
Db	61	HRARCEVGRVLA VVNDPRGPFVGLIACVOLERVLETAASAAIFERRGPALSRERLLYL	120

```
QY 121 ITNYLPSVSLTKRGDEVPDRTLFAHVALCAIGRLGTIVTYDTSLDAAIAPFRHLDP 180
DB 121 ITNYLPSVSLTKRGDEVPDRTLFAHVALCAIGRLGTIVTYDTSLDAAIAPFRHLDP 180
QY 181 ATREGVRRRAAEALAGRTWAPGVEALHTLLSTAVNNMLRDRWSLVAERRRQAGIA 240
DB 181 ATREGVRRRAAEALAGRTWAPGVEALHTLLSTAVNNMLRDRWSLVAERRRQAGIA 240
QY 241 GHTYLOASEKFKIWGAESAPAPERGYKTGAPGAMDTSPAASVPAPQVAVRARQVAS 296
DB 241 GHTYLOASEKFKIWGAESAPAPERGYKTGAPGAMDTSPAASVPAPQVAVRARQVAS 296
QY 297 SSSSSFPAPADMNVPVSAGAPAPPPGSGSYLWIPASHYNQLVTGOSAPRHPPLTACGL 356
DB 297 SSSSSFPAPADMNVPVSAGAPAPPPGSGSYLWIPASHYNQLVTGOSAPRHPPLTACGL 356
QY 301 S-----PVLPPMNPVPTSGTPAPAPPGDGSYLWIPASHYNQLVAGHAAPQPHSAFGF 354
DB 301 S-----PVLPPMNPVPTSGTPAPAPPGDGSYLWIPASHYNQLVAGHAAPQPHSAFGF 354
QY 357 P-AAGTVAYGHPGAGSPHYPPPAHPYPGMLFAGSPLEAQIAALVGAIAAADRQAGLP 415
DB 357 P-AAGTVAYGHPGAGSPHYPPPAHPYPGMLFAGSPLEAQIAALVGAIAAADRQAGLP 415
QY 355 PAAGSVAYGHPGAGLSQHPYHVAHQTPGVLFSGSPLEAQIAALVGAIAAADRQAGGP 414
DB 355 PAAGSVAYGHPGAGLSQHPYHVAHQTPGVLFSGSPLEAQIAALVGAIAAADRQAGGP 414
QY 416 AAAGDHGIRGSAKRREHVEQPEYDCGRDEPDRFPYYPGEARPEPRPVDSRRAARQASG 475
DB 416 AAAGDHGIRGSAKRREHVEQPEYDCGRDEPDRFPYYPGEARPEPRPVDSRRAARQASG 475
QY 476 PHETITLVGAVTSLQOELAHMRARTHAPYGPYPVGPYHHPHADTE-TPAQPPRYPAKA 534
DB 476 PHETITLVGAVTSLQOELAHMRARTHAPYGPYPVGPYHHPHADTE-TPAQPPRYPAKA 534
QY 535 VYLPHPHIAAPPGLPGA--VPPPSYPVAVTPGPAPPLHQPSPAHAHPPPPPGPTPPP 592
DB 535 VYLPHPHIAAPPGLPGA--VPPPSYPVAVTPGPAPPLHQPSPAHAHPPPPPGPTPPP 592
QY 534 VYRPPHAGYPPGPGPASHAPTTPYAAACPPGPPP---PCPSTQTRAPLPTTEPAFTP 590
DB 534 VYRPPHAGYPPGPGPASHAPTTPYAAACPPGPPP---PCPSTQTRAPLPTTEPAFTP 590
QY 593 AASLPQEPAGAEAGALVNASSAAHVNDTARAADLFVSQMMGSR 637
DB 593 AASLPQEPAGAEAGALVNASSAAHVNDTARAADLFVSQMMGSR 637
RESULT 2
US-10-214-932-86
; Sequence 86, Application US/10214932
; Publication No. US2003010070A1
; GENERAL INFORMATION:
; APPLICANT: HWANG, Inhwan
; APPLICANT: KIM, Dae Heon
; APPLICANT: LEE, Yong Jik
; TITLE OF INVENTION: SYSTEM FOR DETECTING PROTEASE
; FILE REFERENCE: APO2/US
; CURRENT APPLICATION NUMBER: US/10/214,932
; CURRENT FILING DATE: 2002-08-08
; NUMBER OF SEQ ID NOS: 133
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 86
; TYPE: PRT
; ORGANISM: Human herpesvirus 1
US-10-214-932-86
Query Match 68.9%; Score 2333; DB 14; Length 635;
Best Local Similarity 74.0%; Pred. No. 4.2e-135;
Matches 477; Conservative 30; Mismatches 120; Indels 18; Gaps 7;
```

```
QY 1 MASAENRELEAPLPDRAPYIVVAGFLALYDSDGPGELALDPTVRAALPPENPLINVD 60
DB 1 MAAADAPGRMEEPPLPDRAPYIVVAGFLALYDSDGPGELALDPTVRAALPPENPLINVD 60
QY 61 HRARCEVGRVAVNDPRGFFVGLIACVOLSERVLETAASAIFERRGPAISREERLLYL 120
DB 61 HRARCEVGRVAVNDPRGFFVGLIACVOLSERVLETAASAIFERRGPAISREERLLYL 120
QY 121 ITNYLPSVSLTKRGDEVPDRTLFAHVALCAIGRLGTIVTYDTSLDAAIAPFRHLDP 180
DB 121 ITNYLPSVSLTKRGDEVPDRTLFAHVALCAIGRLGTIVTYDTSLDAAIAPFRHLDP 180
```

```
QY 181 ATREGVRRRAAEALAGRTWAPGVEALHTLLSTAVNNMLRDRWSLVAERRRQAGIA 240
DB 181 ASREGARRLAAEALASLGRTPWPGVEALHTLLSTAVNNMLRDRWSLVAERRRQAGIA 240
QY 241 GHTYLOASEKFKIWGAESAPAPERGYKTGAPGAMDTSPAASVPAPQ---VAVRARQVAS 296
DB 241 GHTYLOASEKFKIWGAESAPAPERGYKTGAPGAMDTSPAASVPAPQ---VAVRARQVAS 296
QY 297 SSSSSFPAPADMNVPVSAGAPAPPPGSGSYLWIPASHYNQLVTGOSAPRHPPLTACGL 356
DB 301 S-----PVLPPMNPVPTSGTPAPAPPGDGSYLWIPASHYNQLVAGHAAPQPHSAFGF 354
QY 357 P-AAGTVAYGHPGAGSPHYPPPAHPYPGMLFAGSPLEAQIAALVGAIAAADRQAGLP 415
DB 355 PAAGSVAYGHPGAGLSQHPYHVAHQTPGVLFSGSPLEAQIAALVGAIAAADRQAGGP 414
QY 416 AAAGDHGIRGSAKRREHVEQPEYDCGRDEPDRFPYYPGEARPEPRPVDSRRAARQASG 475
DB 415 -AAGDPGVGRSGKRRRYEAGPSESICYDQDEPDADYPYYPGEARGAPRGVDSRRAARHSPG 473
QY 476 PHETITLVGAVTSLQOELAHMRARTHAPYGPYPVGPYHHPHADTE-TPAQPPRYPAKA 534
DB 474 TNETITLVGAVTSLQOELAHMRARTSAPYGYMTPVAHYRQVGPBEPPTTTHPALCPPEA 533
QY 535 VYLPHPHIAAPPGLPGA--VPPPSYPVAVTPGPAPPLHQPSPAHAHPPPPPGPTPPP 592
DB 534 VYRPPHAGYPPGPGPASHAPTTPYAAACPPGPPP---PCPSTQTRAPLPTTEPAFTP 590
QY 593 AASLPQEPAGAEAGALVNASSAAHVNDTARAADLFVSQMMGSR 637
DB 591 AATGSGQPEASNAEAGALVNASSAAHVNDTARAADLFVSQMMGAR 635
RESULT 3
US-10-779-597-30
; Sequence 30, Application US/10779597
; Publication No. US20040234953A1
; GENERAL INFORMATION:
; APPLICANT: Oregon Health & Science University
; APPLICANT: Wong, Scott W.
; APPLICANT: Axthelm, Michael K.
; APPLICANT: Hansen, Scott G.
; TITLE OF INVENTION: JAPANESE MACAQUE HERPESVIRUS NUCLEIC ACID SEQUENCES AND THEIR USE
; FILE REFERENCE: 178-67426
; CURRENT APPLICATION NUMBER: US/10/779,597
; CURRENT FILING DATE: 2004-02-12
; PRIOR APPLICATION NUMBER: 10/276,524
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: PCT/US01/16274
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: 60/205,652
; PRIOR FILING DATE: 2000-05-18
; NUMBER OF SEQ ID NOS: 172
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 30
; LENGTH: 536
; TYPE: PRT
; ORGANISM: Japanese Macaque Herpesvirus
US-10-779-597-30
Query Match 10.4%; Score 351; DB 17; Length 536;
Best Local Similarity 23.3%; Pred. No. 1.8e-13;
Matches 143; Conservative 78; Mismatches 249; Indels 144; Gaps 22;
```

```
QY 21 IYVAGFLALYDSDG-PGELALDPTVRAALPPENPLINVDHRCARCEVGRVAVNDPRG 79
DB 4 VYVGVVDVWSLPKIEKDIYLEPSIVATLLPTNPLPINIEHVPEAHVGHVGLFQVTHG 63
QY 80 PFVGLIACVOL-----ERVLETAASAIFERRGPAISREERLLYLITVLPVSLSTKR 135
DB 64 IFCLKGLTSPDFLALASRLAGDSRAAQIQINMP-----RDPLEMLHTWLTPELSLSLHP 119
QY 136 GDEVPDR-TLFAHVALCAIGRLGTIVTYDTSLDAAIAPFRHLDPATREGVRRRAEAE 194
```

Db 120 EELQDPNHPAPAFQVSVLCALGRRGSIAYVGPDTWVWSK---DSUTRDEAGKITSKCL 176  
Qy 195 LALAGRTWAPGVEALHTLLSTAVNNMLDRWSLVAERROQAGIAGHTYLOASEKEKIW 254  
Db 177 DLCEQVTPPEFAFLETLMAKADAGFIRDRTDLLTKDGVARVARTYLKASQ----- 231  
Qy 255 GAESAPAPERGYKTGAPGAMDTSPAASVPAPQAVRARQVASSSSSSFFAPADMPVSA 314  
Db 232 ----SPSSQHG-----GNRDTQMSALPDNDITI-PKSTFTLWTVQSSLDHMRNQGRAY 280  
Qy 315 SGAPAPPPGCGSY-LMIPASHYNQLVTGQSAHPHPLTACGLPAAGTVAYGHGCGPSP 373  
Db 281 VSAP-PSMPATAAYPSWIPP-----PELTV-----P 305  
Qy 374 HYPPPPAHPYP-CMLFAGSPSLQAIAALVGAIAADRQAGGLPAAAGDHGIRGSAKRER 431  
Db 306 SYAPPVAPPPFQSAFAPQSPYATYS-----PTYGYAQASRH-----QKRK 350  
Qy 432 HEVEQBYDCGRDEPDRDPFYPGEARPEPRPVDSSRAARQASGPHETITLALVGAVTSLQ 491  
Db 351 RDVEL-----SDEPVFPG-----EVGIHKDVMALSKNILDIO 383  
Qy 492 QELAHMARATHAPGYPPVPGYHHPHADTETPAOPP-RYPKAVYLPHPHIAPPGL 549  
Db 384 ADLRDLKRAASQTSG-----EKDTRDQRPQPPVQVSWPQTYASAPYL----- 425  
Qy 550 SGAVPPPSYPVAVTPGAPPLHQPS-----AAHAPPPPPPGPTPPPAASLPQPEAP 602  
Db 426 --AYQPOWY-----PGDTHLHASQBYQIAQGIQQTQPPPPQASHHAGLATQFVAPAP 477  
Qy 603 GAEGALVNASAA 616  
Db 478 AAQESAMNAVPSA 491

## RESULT 4

US-10-437-963-159995  
; Sequence 159995, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 159995  
; LENGTH: 534  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(534)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_59316C.1.pap  
US-10-437-963-159995

Query Match 8.7%; Score 294.5; DB 16; Length 534;  
Best Local Similarity 29.3%; Pred. No. 5.2e-10;  
Matches 119; Conservative 19; Mismatches 137; Indels 131; Gaps 18;  
Qy 256 AESAPAPERGYKTGAPGAMDTSPAASVPAPQAVRARQVASSSSS-SSSFAP----- 306

Db 139 AAAAAAPPAR-----APPATSPPPPPPPPPPPPPPPSSSSSSCSATSPPPCWAAQGT 193  
Qy 307 ADMNPVSASGAPAPPPPGDGSYLWIPASHYNQLVTGQSAHPHPLTAC----- 354  
Db 194 AALPPVSSSSXAPPPPPPP-----PXS-----PAPPQPPPPACTRTTSTXPDERR 239  
Qy 355 -----GLPAAGTVAYGHGAG-----PSPHY-----PPPAHPYPMGLFAGSPPL 394  
Db 240 EABQPRPRPASA-----PGVNHDRVPPPPHRELMPFPXPPPPPPPPPPPPPPETEOPPP 293  
Qy 395 EAQIA-----ALVGAIAADRQAGGLPAAAGDHGIRGSAKRERHEVEQPEYDCG 442  
Db 294 SAYLSPSHSGYSSAPPLPQATAAARAPFRPALPH---CGCA---RQHERLRP----- 344  
Qy 443 RDEPDRDPFYPGEARPEPRPVDSSRAARQASGPHETITLALVGAVTSLQELAHMARATH 502  
Db 345 -----PRXPPPPSPSPAAPVPAASP-----SVLPAPPAS 372  
Qy 503 APYGP-----YPPVGPYHHPHADTETPAOPPYPYPAKAVYLPHPHIAPPGLSGAVPPPSY 558  
Db 373 PPEXPPPVLPPLPPPCPSPLPAPLRSSPSPPTXPAALHTLP-----PPTPCAGKPPPPX 427  
Qy 559 PPVAVTPGAPPLHQ-----PSAHAHAPPPPPPGPTPPPAASLPQPEAP 602  
Db 428 PPPPPPPXPRPADSUPPPSPSAPPSPPP-PPPPPPPLPPPPRAP 472

## RESULT 5

US-10-369-493-3962  
; Sequence 3962, Application US/10369493  
; Publication No. US20030233675A1  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Chen, Xianfeng  
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
; FILE REFERENCE: 38-10(52052)B  
; CURRENT APPLICATION NUMBER: US/10/369,493  
; CURRENT FILING DATE: 2003-02-28  
; PRIOR APPLICATION NUMBER: US 60/360,039  
; PRIOR FILING DATE: 2002-02-21  
; NUMBER OF SEQ ID NOS: 47374  
; SEQ ID NO 3962  
; LENGTH: 639  
; TYPE: PRT  
; ORGANISM: Neurospora crassa  
US-10-369-493-3962

Query Match 8.6%; Score 292; DB 14; Length 639;  
Best Local Similarity 24.0%; Pred. No. 9e-10;  
Matches 126; Conservative 28; Mismatches 190; Indels 180; Gaps 20;  
Qy 176 RHLDP-ATREGVRRAAEAELALAGRT-----WAPGVEALT-----HTLLST 216  
Db 177 KHASPTATPPESPRAAAPACRMWAGSTNGINGYKYPSEFATLDAFDLWREHFDDDLKQ 236  
Qy 217 AVNNMLRDRWSLVAE---RRRQAGIAGHTYLOASEKFIWGAESAPAPERGYKTGAPCAM 274  
Db 237 GLDDDFIRENQDFIIEFLKQEQAKQGIHT-----STPPP-----PPPPPSAP 278  
Qy 275 DTSFAASVPAPQAVRARQVASSSSSSFFAPADMPVSAAGAPAPPPGSGVLTWPAS 334  
Db 279 PNGFAMRAPPPPPPAAPRSVSES-----ITPSTSRRGVPPPPPP-----PAR 321  
Qy 335 HYNQLVT-GOSAPRHPPLTACGLPAAGTVAYGH---PGAGSPHYPPPPRAHP----- 382  
Db 322 RSKLDTENHQBAPPPRFAVPPPIADAGKFAHSDPPRHTPSAPGPPPPPPKPTPLEDQ 381  
Qy 383 YPGMLFAGSPPLEAQIAALVGAIAADRQAGGLPAAAGDHGIRGSAKRERHEVEQPEYDCG 442

Db 382 DPSQRESVPPFTGQ----- 396  
 QY 443 RDEPDRDPFYGCEARPERPRVDSPRAARQAGSPHETITIALVAVTSLQOELAHMRATH 502  
 Db 397 -----RSVFPFPPSRSSVPPPPRNSAAQPPLP-----PKAP 429  
 QY 503 APYGVPPVGVYHHHADTETAQPRRYPAKAVYLPPLP-----HIAPGPPILSGAVPPPS 557  
 Db 430 GRAPPLFPASSRPPMLPTRTSFAPPOAPPLPTSNAPPPPLPATAOPPPLPATSAPPP 489  
 QY 558 YPFVAVTPGAPPLP---HQFSP-----AHAPPPPPPPG-----PTPPPA 593  
 Db 490 PFPAPPAP-PAPPLPAHAHAPPPPPMPMPAPSGGAPPPPPPPPPGGMGVPPPPPPPP 548  
 QY 594 ASLPQPEAP-----GAECALVNASSAHV-----NVDTARAAD 627  
 Db 549 GGMPPPPAPAPFPVDGSRSAVLGDKTAGGIRALKKVDRSQIRD 592

RESULT 6

US-10-408-765A-2285  
 ; Sequence 2285, Application US/10408765A  
 ; Publication No. US20040101874A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Ghosh, Soumitra S.  
 ; APPLICANT: Fahy, Eoin D.  
 ; APPLICANT: Zhang, Bing  
 ; APPLICANT: Gibson, Bradford W.  
 ; APPLICANT: Taylor, Steven W.  
 ; APPLICANT: Glenn, Gary M.  
 ; APPLICANT: Warnock, Dale B.  
 ; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION  
 ; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME  
 ; FILE REFERENCE: 660088.465  
 ; CURRENT APPLICATION NUMBER: US/10/408,765A  
 ; CURRENT FILING DATE: 2003-04-04  
 ; NUMBER OF SEQ ID NOS: 3077  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 2285  
 ; LENGTH: 825  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-10-408-765A-2285

Query Match 8.4%; Score 284; DB 16; Length 825;  
 Best Local Similarity 24.7%; Pred. No. 3.7e-09;  
 Matches 159; Conservative 58; Mismatches 230; Indels 196; Gaps 30;

QY 35 PGELALDPTVRAALPPENPLPINVDHRARCEVGRVIAV-----VNDPRGPF 81  
 Db 185 PG-LGLDPNK-----PPELPSTVNAEPLGLIQSGPHQAAPPPPPPPPPAPASEPKGL 238  
 QY 82 FVGLIACVOLERVLETAASAAIFERRGPALSREERLLYLTNVLPSVLSLTKRRGDEVPP 141  
 Db 239 TSPIFCSTPKLLKT-SFHLRLRRDPFPQPKL-----YAOEYFEADEDKADVPA 291  
 QY 142 DRTLFAHVALCAIGRLGTVTYDTSLSDAIAAPFRHL-----PAT 182  
 Db 292 DirlNP-----RLPLVSSCRS-RPALSPGLGIDFCPPNPGPDGPRRRKRPTAK 342  
 QY 183 REGVREAAEAL-ALAGTWPAGVEALHTLLSTAVNNMLRDRWSLVAERRRONGIAG 241  
 Db 343 RDGPPRGRPRIRPLEVTTAGPASASTPT---DGAKKPRGR-----GGRKXABEAG 394  
 QY 242 HTYLOASEKFKWGAESAPAPERGYKTGAPGAMD-----TSPAASVPAPQVAVRAOVASS 297  
 Db 395 GTFLEPLKPKLT-KLSVPKAGEGLTSGDALISGTDHNSLDSLTREKIEAKIKEY--- 449  
 QY 298 SSSSSPPADNMNIVSASGAPAPPPPGDGSYLWIPASHYNQLVTGQAPRHPPLTACGLP 357  
 Db 450 -----EEXQPEKMG-----FMASFLDFLKSCK----RHPPLYOAGLT 483  
 QY 358 AAGTVAGHGAGPSHYPP-----PPAHPPYGMFLFAGPSPLEAQIAALVGAI-A 406

Db 484 -----PPLSPKSVPPSPVARGLQOPATP-----AVPHPPSGAGLGGALEA 528  
 QY 407 ADROQAGGLPAAAGDHGIRGSAKRRRHEVEQPEYDCGRDEPDRDPFYYPGEARPERPVD 466  
 Db 529 ABSEGLGL-----GCPSPCKRLDELK-----RNLETLSPFSFSEEDSVAK 569  
 QY 467 RRAARQASGPHETITIALVAVTSLQOELAHMRATHAPYGPYPV-GPVYHHHADTETA 525  
 Db 570 NRDLQ-----ESTSSAISALDD-----PPLAGP-----KDTSTPD 599  
 QY 526 QPPRYPAKAVYLPPL-----HIAPGPPILSGAVPPSPVAVTPGPAPPLHQPSA 577  
 Db 600 GPFLAPAAAVCPGPPPLGFLPSANSNGTEPPELLEKPPPT-PPAPTQOPPPPPPPPP 658  
 QY 578 HAHF-----PPPPPGTTPPPAASLPOPEAPGAEGA 608  
 Db 659 PALPSPPPPLVAPTSPSPPPPLPPPPPPA-MSPSPPPPPPPAAA 700

RESULT 7

US-10-437-963-191498  
 ; Sequence 191498, Application US/10437963  
 ; Publication No. US20040123343A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: La Rosa, Thomas J.  
 ; APPLICANT: Kovalic, David K.  
 ; APPLICANT: Zhou, Yihua  
 ; APPLICANT: Cao, Yongwei  
 ; APPLICANT: Wu, Wei  
 ; APPLICANT: Boukharov, Andrey A.  
 ; APPLICANT: Barbazuk, Brad  
 ; APPLICANT: Li, Ping  
 ; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
 ; FILE REFERENCE: 38-21(53221)B  
 ; CURRENT APPLICATION NUMBER: US/10/437,963  
 ; CURRENT FILING DATE: 2003-05-14  
 ; NUMBER OF SEQ ID NOS: 204966  
 ; SEQ ID NO 191498  
 ; LENGTH: 483  
 ; TYPE: PRT  
 ; ORGANISM: Oryza sativa  
 ; FEATURE:  
 ; NAME/KEY: unsure  
 ; LOCATION: (1)..(483)  
 ; OTHER INFORMATION: unsure at all xaa locations  
 ; FEATURE:  
 ; OTHER INFORMATION: Clone ID: PAT\_MBT4530\_87810C.1.psp  
 US-10-437-963-191498

Query Match 8.3%; Score 281.5; DB 16; Length 483;  
 Best Local Similarity 28.1%; Pred. No. 2.9e-09;  
 Matches 99; Conservative 12; Mismatches 130; Indels 111; Gaps 15;

QY 303 FPAPADMNIVSASGAPAPPPPGDGSYLWIPASHYNQLVTGQAPRHPPTACGLPAAGTV 362  
 Db 85 YPPPX---PPYPGXYPYPPPPYPYPPYPPXK---RPTSDPRPRQPP-PCPPEXPPPP 136  
 QY 363 AVGHGAGPSPH-----YPPPPAHPPYGMFLFAGPSPLEAQIAALV 402  
 Db 137 FXPPPPPPPPYPPPPPPPPWGXYPYPPYPLPYPPXSPYFGWYPPYP----- 187  
 QY 403 GAIAADROAGGLPAAAGDHGIRGSAKRRRHEVEQPEYDCGRDEPDRDPFYYP----- 454  
 Db 188 -----GWP-----YPPPPYPPYAGRYPPYPPYPPYPPYPPXDIH 221  
 QY 455 -----GEARPERPRVDSRRARQASGPHETITIALVAVTSLQOELAHMRATHAPYGPYP 511  
 Db 222 LFTSDPRPPPPPPRRPXAAPPAPPPGPP-----PXTSPDPLVGRASPALP---PPPPPP 273  
 QY 512 GPY-----HHPHADTETPAOPRYPAKAVYLPPLPPHPIAPPGGPPLSGAVP- 554



RESULT 9

US-10-437-963-142229

Sequence 142229, Application US/10437963

Publication No. US20040123343A1

GENERAL INFORMATION:

APPLICANT: La Rosa, Thomas J.

APPLICANT: Kovalic, David K.

APPLICANT: Zhou, Yihua

APPLICANT: Cao, Yongwei

APPLICANT: Wu, Wei

APPLICANT: Boukharov, Brad

APPLICANT: Barbazuk, Brad

APPLICANT: Li, Ping

TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with

FILE REFERENCE: 38-21(53221)B

CURRENT APPLICATION NUMBER: US/10/437,963

CURRENT FILING DATE: 2003-05-14

NUMBER OF SEQ ID NOS: 204966

SEQ ID NO 142229

LENGTH: 426

TYPE: PRT

ORGANISM: Oryza sativa

FEATURE:

NAME/KEY: unsure

LOCATION: (1)..(426)

OTHER INFORMATION: unsure at all Xaa locations

FEATURE:

OTHER INFORMATION: Clone ID: PAT-MRT4530\_43255C.1.pap

US-10-437-963-142229

Query Match 8.1%; Score 275.5; DB 16; Length 426;

Best Local Similarity 27.9%; Pred. No. 5.9e-09;

Matches 107; Conservative 26; Mismatches 185; Indels 65; Gaps 12;

Qy 258 SAPAPERGYKTGAPGMDTSPAAS-----VPAPQVAVRARQVASSSSSSFFAPADMNV 312

Db 60 AATGPGRGPDPAAPTVAADPTTTPSLPPLPPSPPLDLAFAADLASSPPPPHVRT 119

Qy 313 SASGAPAPPPGDCGSLYLPASHYNQLVTGQAPRHPPLTACGLPAAGTVAYGH---PG 368

Db 120 SPLSSPAPPPP---SPVIPPQRPXPASPPAPGAPRQPRPTVVGCTSMRHGCLSP 176

Qy 369 AGPSPHYPPPAHPYPCGMLFAGSPGLEAQIALVGAIAADROAGGLPAAAGDHGIRSAK 428

Db 177 AEPSCSXPPPPAVAPPQPTRATPAPQAPLXA-----AAGSSRGPDPAHTAAVAAPTAT 231

Qy 429 RRRHEVQPEVDCGRDEPDRDF-----PYYPGEAR-PEPRPVDSSRAARQASGPHETI 480

Db 232 PPLPPLPPSPPLDLAAPTTLASLSQFPFPPTKKKPPSPXPVPPRHVTPPPPRPH--- 288

Qy 481 TALVGAVTSLQQELAHMRARTHAPYGPYPVGVYHHPHADTETPAQPPRYPAKAVLPP- 539

Db 289 -----VPSIPP-----PPXPFPPLPPLPWPPPLLSPAWXFPSPSTPPARLPPX 332

Qy 540 PHIAPPGPPSLGAV-PPPSYPPVAVTGPAPPLHQPSPAHAPPP-PPPGGTP----- 590

Db 333 PSISRPPPPRTPLPDPPSPPPPLFFPPPPPPXSFVNRHSPQPPHSPSPSPSLPLP 392

Qy 591 -----PPAASLPQPEAP 602

Db 393 LVAPPHRRRSSPPAVRRSSPPAP 415

RESULT 10

US-10-437-963-126337

Sequence 126337, Application US/10437963

Publication No. US20040123343A1

GENERAL INFORMATION:

APPLICANT: La Rosa, Thomas J.

APPLICANT: Kovalic, David K.

```
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 126337
; LENGTH: 506
; TYPE: PRT
; ORGANISM: Oryza sativa
; NAME/KEY: unsure
; LOCATION: (1)..(506)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_28894C.1.pep
US-10-437-963-126337

Query Match      8.1%; Score 273; DB 16; Length 506;
Best Local Similarity 29.2%; Pred. No. 1e-08;
Matches 109; Conservative 18; Mismatches 130; Indels 116; Gaps 19;

QY 270 AFGAMDTSPAASVPAQVAVRA---RQVASS-----SSSSFFPAPADMN-----P 311
DQ 30 SPIAVIASPVRAPSLARVPACRRRVASSHPAALCVLSPSPVPRXPSSPPPPAP 89
QY 312 VSASGAPAPPPGDSYLIWIPASHYNQLVTGOSAPRHPLTACGLPAACTVAYGHGAGP 371
DQ 90 XPRACAPPPPPPD-----VV-----PPPSAPXPPAGPPARRPP---P 125
QY 372 SHYP-----PPPA---HPYRCMLFAGSPLEAQIAALVGAIAADROAGGLPAAAG 419
DQ 126 SPLRFDLGGREGGEPFPAASRRPPAP---ASHPHPPAR-----SGG 163
QY 420 DHGIRGSAKRHEVEQDYDCGRDEPDRDFYYPGEARPEP-RP-VDSRRAARQASGP- 476
DQ 164 XEGEKPATPRTPXRFSPDAVRSPDGIRSPSAGLPHLRPDLEGREGGEPXP 223
QY 477 --HETITALVAVTSLQELAHMRARTHAPYGPVPGVYHHP-----HADTETAPQPR 529
DQ 224 CARPPVRLPDVR-----PSPASPLPPPARSGGEGXEPXP 265
QY 530 YPAKAVYLPPIHAPGPPPLSGNVPPSPYPPVAVTGPAPPLHQPSAHAPPPPPGPT 589
DQ 266 PPDQ-----PPRLAPQRPD--PAVPPP--PPREGPPPPPPPPPPPP-----PPPPPPPPPL 313
QY 590 PPPAASLPOEAP 602
DQ 314 PPPLAPRCEPPLP 326

RESULT 11
US-10-437-963-179715
; Sequence 179715, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
```

```
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 179715
; LENGTH: 521
; TYPE: PRT
; ORGANISM: Oryza sativa
; NAME/KEY: unsure
; LOCATION: (1)..(521)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_7714C.1.pep
US-10-437-963-179715

Query Match      8.0%; Score 271.5; DB 16; Length 521;
Best Local Similarity 27.1%; Pred. No. 1.3e-08;
Matches 95; Conservative 20; Mismatches 117; Indels 119; Gaps 17;

QY 256 AESAPAPRGYKTGAPGAMDTSPAASVPAQVAVRAQVASSSSSSSPPA-PADMNPSA 314
DQ 20 AATSPSP-----PPGRAGNVPTSAREPSVT-----PTWPTTTPGPTP 57
QY 315 SGAPAPPPPGDSY--LWIPASHYNQLVTGOSAPRH---PPLTACGLPAACTVAYGHGPA 369
DQ 58 SARPLPLPPTTSAHPEAATPPPHAPPAPQDPXPHSRPPPLTLTVPP-----HRRH 110
QY 370 GPSPHYPPPAHPYPMGLFAGSPLEAQIAALVGAIAADROAGGLPAAAGDHGIRGSAKR 429
DQ 111 YENPHLPPPPQPLP---YTPP-----RRSPPR 135
QY 430 RHVEVEQDYDCGRDEPDRDFYYPGEARPEPVPDSRAARQASGPHETITALVAVT 489
DQ 136 SRPPTTP-----SRPPTTPAHPPPPQPTP---RR--HHPTTTPH---APAAPT 179
QY 490 LQELAHMRARTHAPYGPVPGVYHHPHADTETPAQPRYPKAVYLPPIHAPGPPPL 549
DQ 180 XPQ-----RTTTP--PPPP-----PPPDTSRTPPTTPPRAPPTPPRPGPPPL 224
QY 550 SGAVPPSPYPPVAVTGPAPPLHQPSAHAPPPPPPPGPTPPPAASLPOE 600
DQ 225 ----PPPPPPAPPPPG-----PPPPPPXPPQPPPTSLPPKE 257

RESULT 12
US-10-437-963-195106
; Sequence 195106, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 195106
; LENGTH: 606
; TYPE: PRT
; ORGANISM: Oryza sativa
; NAME/KEY: unsure
; LOCATION: (1)..(606)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_91087C.1.pep
```

US-10-437-963-195106

Query Match 8.0%; Score 270; DB 16; Length 606;  
Best Local Similarity 27.1%; Pred. No. 1.9e-08;  
Matches 110; Conservative 19; Mismatches 143; Indels 13

Qy	260	PAPERGYKTGAPGAMDTSPAA	-----VPAPQAVARQAVASSSSSSSFFPAPA	307
Db	196	PAPHPNGMPTGSP	---TISPAAEPPPLPKVPXPTPVPPPLPRXISSSSPTPPSAPL	252
Qy	308	DMNPVSASGAPAPPPPGDGSYLWTPASHYN	-----QLVTQSGAPRHPPLTACGLPA	358
Db	253	-----SSPLPCPPPP	---IYCPHKLYRFPAPRLAPPSLV	297
Qy	359	AGTVAYGHGAGP	-----SPHYPP	389
Db	298	HGNLPLSPPAKAPGNKPPPPVTPRPLPGSPXPPPPVPSRPLPVDXPPXPPPSILRA	357	
Qy	390	GPSPLEAQIAALVCAIADNDROAGGLPAAAGDHGIRGSAKRHRHEVEOPEYDCGRDEPRD	449	
Db	358	GPOFXTPTPPSPAALAIACRPPPSRPAAT	-----TAARRRPPSLPTAADATRPppppp	410
Qy	450	FPVYPGARDEPRVDSSRAARQSGPHEITALTALVGAVTSLQELAHMRARTHAPYGPY	509	
Db	411	PPHSXPSLSLPEDEKTPPPPPPPSP	-----PPLRURACPTTPPH	452
Qy	510	PVGPYHHPHADTETPAOPPRYPAKAVLPLPHPIAPP	-GPPLSGAVPPSPYPPVAVTPGPA	568
Db	453	XPPPY	-----PSPVSP	489
Qy	569	PPLH	-----QPSGAHNP	597
Db	490	PPHPLPLSLGISFPPTXMPSPRPPPSLSPPPXPAKXPPPPAGL	535	

**RESULT 13**

US-10-437-963-149274  
; Sequence 149274, Application US/10437963  
; Publication No. US20040123343A1

**GENERAL INFORMATION:**

APPLICANT: La Rosa, Thomas J.

APPLICANT: Kovalic, David K.

APPLICANT: Zhou, Yihua

APPLICANT: Cao, Yongwei

APPLICANT: Wu, Wei

APPLICANT: Boukharov. Andrei.

APPLICANT: DOBNAKOV, ANDREY  
: APPLICANT: Barbazuk, Brad

APPLICANT: BALDZUK, Blau  
APPLICANT: I. J. Ping

APPLICANT: LI, FENG  
TITLE OF INVENTION: Pico Nuc

;	TITLE OF INVENTION: RICE NUCLEOTIDES
1	INVENTION: Plants and

1. TITLE OF INVENTION: Plants and

; FILE REFERENCE: 38-21(53221)B

;	CURRENT APPLICATION NUMBER:	0
;	CURRENT FIVE YEAR DATE:	00000000

2003-05

; NUMBER OF SEQ ID NOS: 204966

; SEQ ID NO 149274

; LENGTH: 373

; TYPE: PRT

; ORGANISM: Oryza sativa

; FEATURE:

```

; NAME/KEY: unsure

```

; LOCATION: (1) .. (373) .

OTHER INFORMATION: unsure at

;; FEATURE:

OTHER INFORMATION: Clone ID:

US-10-437-963-149274

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

Query Match 7.9%:

Query Machine	Best Local Similarity	Best Local Similarity
7.5%	26.1%	

BEST LOCAL SIMILARITY 20.1%;  
MATCHES 96. Conservative

WATCIN 36; COMBETVATIVE

Q 257 P 561 P 562 P 563 P 564 P 565 P 566 P 567 P 568 P 569 P 570 P 571 P 572 P 573 P 574 P 575 P 576 P 577 P 578 P 579 P 580 P 581 P 582 P 583 P 584 P 585 P 586 P 587 P 588 P 589 P 590 P 591 P 592 P 593 P 594 P 595 P 596 P 597 P 598 P 599 P 600 P 601 P 602 P 603 P 604 P 605 P 606 P 607 P 608 P 609 P 610 P 611 P 612 P 613 P 614 P 615 P 616 P 617 P 618 P 619 P 620 P 621 P 622 P 623 P 624 P 625 P 626 P 627 P 628 P 629 P 630 P 631 P 632 P 633 P 634 P 635 P 636 P 637 P 638 P 639 P 640 P 641 P 642 P 643 P 644 P 645 P 646 P 647 P 648 P 649 P 650 P 651 P 652 P 653 P 654 P 655 P 656 P 657 P 658 P 659 P 660 P 661 P 662 P 663 P 664 P 665 P 666 P 667 P 668 P 669 P 670 P 671 P 672 P 673 P 674 P 675 P 676 P 677 P 678 P 679 P 680 P 681 P 682 P 683 P 684 P 685 P 686 P 687 P 688 P 689 P 690 P 691 P 692 P 693 P 694 P 695 P 696 P 697 P 698 P 699 P 700 P 701 P 702 P 703 P 704 P 705 P 706 P 707 P 708 P 709 P 710 P 711 P 712 P 713 P 714 P 715 P 716 P 717 P 718 P 719 P 720 P 721 P 722 P 723 P 724 P 725 P 726 P 727 P 728 P 729 P 730 P 731 P 732 P 733 P 734 P 735 P 736 P 737 P 738 P 739 P 740 P 741 P 742 P 743 P 744 P 745 P 746 P 747 P 748 P 749 P 750 P 751 P 752 P 753 P 754 P 755 P 756 P 757 P 758 P 759 P 760 P 761 P 762 P 763 P 764 P 765 P 766 P 767 P 768 P 769 P 770 P 771 P 772 P 773 P 774 P 775 P 776 P 777 P 778 P 779 P 780 P 781 P 782 P 783 P 784 P 785 P 786 P 787 P 788 P 789 P 790 P 791 P 792 P 793 P 794 P 795 P 796 P 797 P 798 P 799 P 800 P 801 P 802 P 803 P 804 P 805 P 806 P 807 P 808 P 809 P 810 P 811 P 812 P 813 P 814 P 815 P 816 P 817 P 818 P 819 P 820 P 821 P 822 P 823 P 824 P 825 P 826 P 827 P 828 P 829 P 830 P 831 P 832 P 833 P 834 P 835 P 836 P 837 P 838 P 839 P 840 P 841 P 842 P 843 P 844 P 845 P 846 P 847 P 848 P 849 P 850 P 851 P 852 P 853 P 854 P 855 P 856 P 857 P 858 P 859 P 860 P 861 P 862 P 863 P 864 P 865 P 866 P 867 P 868 P 869 P 870 P 871 P 872 P 873 P 874 P 875 P 876 P 877 P 878 P 879 P 880 P 881 P 882 P 883 P 884 P 885 P 886 P 887 P 888 P 889 P 890 P 891 P 892 P 893 P 894 P 895 P 896 P 897 P 898 P 899 P 900 P 901 P 902 P 903 P 904 P 905 P 906 P 907 P 908 P 909 P 910 P 911 P 912 P 913 P 914 P 915 P 916 P 917 P 918 P 919 P 920 P 921 P 922 P 923 P 924 P 925 P 926 P 927 P 928 P 929 P 930 P 931 P 932 P 933 P 934 P 935 P 936 P 937 P 938 P 939 P 940 P 941 P 942 P 943 P 944 P 945 P 946 P 947 P 948 P 949 P 950 P 951 P 952 P 953 P 954 P 955 P 956 P 957 P 958 P 959 P 960 P 961 P 962 P 963 P 964 P 965 P 966 P 967 P 968 P 969 P 970 P 971 P 972 P 973 P 974 P 975 P 976 P 977 P 978 P 979 P 980 P 981 P 982 P 983 P 984 P 985 P 986 P 987 P 988 P 989 P 990 P 991 P 992 P 993 P 994 P 995 P 996 P 997 P 998 P 999 P 1000 P 1001 P 1002 P 1003 P 1004 P 1005 P 1006 P 1007 P 1008 P 1009 P 1010 P 1011 P 1012 P 1013 P 1014 P 1015 P 1016 P 1017 P 1018 P 1019 P 1020 P 1021 P 1022 P 1023 P 1024 P 1025 P 1026 P 1027 P 1028 P 1029 P 1030 P 1031 P 1032 P 1033 P 1034 P 1035 P 1036 P 1037 P 1038 P 1039 P 1040 P 1041 P 1042 P 1043 P 1044 P 1045 P 1046 P 1047 P 1048 P 1049 P 1050 P 1051 P 1052 P 1053 P 1054 P 1055 P 1056 P 1057 P 1058 P 1059 P 1060 P 1061 P 1062 P 1063 P 1064 P 1065 P 1066 P 1067 P 1068 P 1069 P 1070 P 1071 P 1072 P 1073 P 1074 P 1075 P 1076 P 1077 P 1078 P 1079 P 1080 P 1081 P 1082 P 1083 P 1084 P 1085 P 1086 P 1087 P 1088 P 1089 P 1090 P 1091 P 1092 P 1093 P 1094 P 1095 P 1096 P 1097 P 1098 P 1099 P 1100 P 1101 P 1102 P 1103 P 1104 P 1105 P 1106 P 1107 P 1108 P 1109 P 1110 P 1111 P 1112 P 1113 P 1114 P 1115 P 1116 P 1117 P 1118 P 1119 P 1120 P 1121 P 1122 P 1123 P 1124 P 1125 P 1126 P 1127 P 1128 P 1129 P 1130 P 1131 P 1132 P 1133 P 1134 P 1135 P 1136 P 1137 P 1138 P 1139 P 1140 P 1141 P 1142 P 1143 P 1144 P 1145 P 1146 P 1147 P 1148 P 1149 P 1150 P 1151 P 1152 P 1153 P 1154 P 1155 P 1156 P 1157 P 1158 P 1159 P 1160 P 1161 P 1162 P 1163 P 1164 P 1165 P 1166 P 1167 P 1168 P 1169 P 1170 P 1171 P 1172 P 1173 P 1174 P 1175 P 1176 P 1177 P 1178 P 1179 P 1180 P 1181 P 1182 P 1183 P 1184 P 1185 P 1186 P 1187 P 1188 P 1189 P 1190 P 1191 P 1192 P 1193 P 1194 P 1195 P 1196 P 1197 P 1198 P 1199 P 1200 P 1201 P 1202 P 1203 P 1204 P 1205 P 1206 P 1207 P 1208 P 1209 P 1210 P 1211 P 1212 P 1213 P 1214 P 1215 P 1216 P 1217 P 1218 P 1219 P 1220 P 1221 P 1222 P 1223 P 1224 P 1225 P 1226 P 1227 P 1228 P 1229 P 1230 P 1231 P 1232 P 1233 P 1234 P 1235 P 1236 P 1237 P 1238 P 1239 P 1240 P 1241 P 1242 P 1243 P 1244 P 1245 P 1246 P 1247 P 1248 P 1249 P 1250 P 1251 P 1252 P 1253 P 1254 P 1255 P 1256 P 1257 P 1258 P 1259 P 1260 P 1261 P 1262 P 1263 P 1264 P 1265 P 1266 P 1267 P 1268 P 1269 P 1270 P 1271 P 1272 P 1273 P 1274 P 1275 P 1276 P 1277 P 1278 P 1279 P 1280 P 1281 P 1282 P 1283 P 1284 P 1285 P 1286 P 1287 P 1288 P 1289 P 1290 P 1291 P 1292 P 1293 P 1294 P 1295 P 1296 P 1297 P 1298 P 1299 P 1300 P 1301 P 1302 P 1303 P 1304 P 1305 P 1306 P 1307 P 1308 P 1309 P 1310 P 1311 P 1312 P 1313 P 1314 P 13

00Y 25/ ESAPAPRGYRTG-----

[illegible]

Db 69 DERPAFPRQGSTSPXPPP

Search completed: January 10, 2005, 19:31:11  
Job time : 150 secs

```
QY 423 IRGSAKRREHVEQPEYDCGRDEPDFFYPGEARPEP-RVVDSSRAARQASGPHEIT 481
Db 471 -----CG-----SQASPPPGTPLASTFSSKSVLPSPSAG 500
QY 482 ALVGAVTSLQQLAH-----MRARTHAPYGPYP---VGPYHHPHADTETPAOPPRYP 531
Db 501 APASAEPLNPELGDSSASEPGLQASQPAESFTQGLVLP-----PAPPP--- 547
QY 532 AKAVYLPPIHAPPGPLSGAVPPPSYPVAVTPGAPPLHQPSPAHAHPPPPPPGCP--- 588
Db 548 -----PPP--LPSPGAVASALPPP-----PGPPPP--PPLFSTGPPPPPPPPPPPLP 589
QY 589 --TPPPAASLPOEPAGAEAGALVNASSA 615
Db 590 NOAPPPP-----PPPPAPPLPASGIFSGSTS 615

RESULT 15
US-10-425-115-231956
; Sequence 231956, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(5322)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 231956
; LENGTH: 280
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(280)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_143137C.1.pap
US-10-425-115-231956

Query Match 7.9%; Score 267; DB 17; Length 280;
Best Local Similarity 27.8%; Pred. No. 1.3e-08;
Matches 84; Conservative 6; Mismatches 84; Indels 128; Gaps 12;

QY 304 PAPADMNPNVSACAPAPPPGDSYLVIPASHYNQLVTCQSAPRHPPLTACGLPAAGTVA 363
Db 65 PPPFHQHRTSPSPPPPPPPPPG-----PPPHFPATPPXPXPP--- 100
QY 364 YGHPGAGSPSPHYPP-PPAHPYFGMLFAGSPSLEAQIALVGAIAADRAQAGLPAAGDHG 422
Db 101 ---PPPPXPXHPPPQPPNP-PPTHPPPTP----- 127
QY 423 IRGSAKRREHVEQPEYDCGRDEPDFFYPGEARPEP-RVVDSSRAARQASGPHEIT 482
Db 128 -----PPRPPPPPPXPTSPPPPPAPSR-----GP----- 152
QY 483 LVGAVTSLQQLAHMRARTHAPYGPYPVGPYHHPHADTETPAOPPRYPKAVYLPPIHI 542
Db 153 -----PPPPXPPNPPPPHP-----TPPPPPPPPPPPPPPPPS 186
QY 543 APPGPLSGAVPPSPVAVTPGAPPLHQPSPAHAHP--PPPPPGTTPPAASLPQPE 600
Db 187 SPPPPPL-----PPP--PPLLPPSPHPPTTTPPPPH-HRQPPPPPPPPPPNPTPPPT 239
QY 601 AP 602
Db 240 PP 241
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 10, 2005, 18:56:04 ; Search time 160 Seconds  
(without alignments)  
1428.191 Million cell updates/sec

Title: US-10-623-429-9

Perfect score: 3385

Sequence: 1 MSAENRERLEAPLPDRAVP.....VNVDTARAADLFVQMMGSR 637

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 segs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A Geneseq\_23Sep04:.\*  
1: Geneseqp1980s:.\*  
2: Geneseqp1990s:.\*  
3: Geneseqp2000s:.\*  
4: Geneseqp2001s:.\*  
5: Geneseqp2002s:.\*  
6: Geneseqp2003as:.\*  
7: Geneseqp2003bs:.\*  
8: Geneseqp2004s:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3385	100.0	637	6	Aae36428 HSV-2 UL2
2	3385	100.0	637	8	Adj77800 Herpes si
3	3370.5	99.6	638	2	Aar71017 HSV-2 pro
4	3359.5	99.2	642	2	Aaw72124 HSV-2 str
5	3351.5	99.0	636	2	Aar64767 HSV-2 pro
6	3267.5	96.5	657	2	Aaw72192 HSV-2 str
7	2333	68.9	635	2	Aar71031 HSV-1 UL2
8	2333	68.9	635	7	Adl18166 Human her
9	2329	68.8	635	2	Aar28634 UL26 prot
10	2326	68.7	635	2	Aar28651 UL26 prot
11	2323.5	68.6	636	2	Aar28636 UL26 prot
12	2322	68.6	635	2	Aar28647 UL26 prot
13	2321	68.6	635	2	Aar28652 UL26 prot
14	2319	68.5	626	2	Aar28640 UL26 prot
15	2319	68.5	635	2	Aar28650 UL26 prot
16	2318.5	68.5	636	2	Aar28637 UL26 prot
17	2318.5	68.5	636	2	Aar28635 UL26 prot
18	2318	68.5	635	2	Aar28649 UL26 prot
19	2306	68.1	635	2	Aar28648 UL26 prot
20	2241	66.2	615	2	Aar28642 UL26 prot
21	2209	65.3	603	2	Aar28641 UL26 prot
22	2017	59.6	514	2	Aar28643 UL26 prot
23	2014.5	59.5	448	2	Aaw72048 HSV-2 str
24	1826	53.9	331	2	Aar64769 HSV-2 ISP
25	1820	53.8	329	6	Aae36429 HSV-2 UL2

26	1432	42.3	350	2	AAW10155	AAW10155 HSV-1 (F)
27	1432	42.3	350	2	AAW09967	AAW09967 HSV-1 (F)
28	1429	42.2	350	2	AAW09960	AAW09960 HSV-1 (F)
29	1429	42.2	350	2	AAW09958	AAW09958 HSV-1 (F)
30	1429	42.2	350	2	AAW09971	AAW09971 HSV-1 (F)
31	1429	42.2	350	2	AAW09973	AAW09973 HSV-1 (F)
32	1426	42.1	350	2	AAW09956	AAW09956 HSV-1 (F)
33	1426	42.1	350	2	AAW09962	AAW09962 HSV-1 (F)
34	1426	42.1	350	2	AAW09975	AAW09975 HSV-1 (F)
35	1426	42.1	350	2	AAW09969	AAW09969 HSV-1 (F)
36	1424	42.1	350	2	AAW09963	AAW09963 HSV-1 (F)
37	1424	42.1	350	2	AAW09985	AAW09985 HSV-1 (F)
38	1424	42.1	350	2	AAW09976	AAW09976 HSV-1 (F)
39	1424	42.1	350	2	AAW09978	AAW09978 HSV-1 (F)
40	1423	42.0	350	2	AAW09959	AAW09959 HSV-1 (F)
41	1423	42.0	350	2	AAW09964	AAW09964 HSV-1 (F)
42	1423	42.0	350	2	AAW09977	AAW09977 HSV-1 (F)
43	1423	42.0	350	2	AAW09972	AAW09972 HSV-1 (F)
44	1422	42.0	350	2	AAW09957	AAW09957 HSV-1 (F)
45	1422	42.0	350	2	AAW09961	AAW09961 HSV-1 (F)

ALIGNMENTS

RESULT 1  
AAE36428  
ID AAE36428 standard; protein; 637 AA.  
XX AAE36428;  
AC AAE36428;  
DT 07-AUG-2003 (first entry)  
XX HSV-2 UL26 full-length antigen.  
DE HSV-2 UL26 full-length antigen.  
XX Herpes simplex virus; HSV; infection; vaccine; therapy; UL26; antigen.  
KW Herpes simplex virus type 2.  
XX OS WO2003020108-A2.  
PN 13-MAR-2003.  
PD 27-AUG-2002; 2002WO-US027341.  
XX PF 04-SEP-2001; 2001US-0317159P.  
XX PR 17-APR-2002; 2002US-0373429P.  
XX (CORI-) CORIXA CORP.  
PA Hosken NA, Day CH;  
XX WPI; 2003-290135/28.  
XX N-PSDB; AAD55178.  
PT Novel isolated polypeptide comprising immunogenic portion of a herpes simplex virus antigen, useful for detecting herpes simplex virus infection in a subject, and for treating the virus infection in a patient.  
PT Claim 1; Page 110-112; 114pp; English.  
XX The invention relates to polypeptides comprising an immunogenic portion of herpes simplex virus (HSV) antigen and to nucleic acid molecules encoding polypeptides. Polypeptides of the invention are useful for detecting and treating HSV infection in a patient. Polynucleotides of the invention are useful as diagnostic reagents for detecting HSV infection in a patient and also as probes or primers. The invention is used to prepare vaccines. The present sequence is HSV-2 strain HG52 UL26 full-length antigen  
Sequence 637 AA;

Query Match		100.0%;	Score 3385;	DB 6;	Length 637;	
Best Local Similarity		100.0%;	Pred. No. 4e-209;			
Matches 637;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
QY	1	MASAEVRERLEAPL	PDRAVPIYVAGFLALYDSGDPGELALD	PDPTVRAALPPENPL	PI	NDV 60
DB	1	MASAEVRERLEAPL	PDRAVPIYVAGFLALYDSGDPGELALD	PDPTVRAALPPENPL	PI	NDV 60
QY	61	HRARCEVRGLAVVND	PRGPFVGLIACVQLSERVLETAASAAI	FERRGPALSRERLLYL	120	
DB	61	HRARCEVRGLAVVND	PRGPFVGLIACVQLSERVLETAASAAI	FERRGPALSRERLLYL	120	
QY	121	ITNYLPSVLSLTKRGDEVP	DRTLFAHVALCAIGRRLG	TIYVDTSLDAAI	APRHLDP	180
DB	121	ITNYLPSVLSLTKRGDEVP	DRTLFAHVALCAIGRRLG	TIYVDTSLDAAI	APRHLDP	180
QY	181	ATREGVRRERAAEAL	LAGRTWAPGVEAL	THLLSTAVNNMLRDRWSI	VAERRRQAGIA	240
DB	181	ATREGVRRERAAEAL	LAGRTWAPGVEAL	THLLSTAVNNMLRDRWSI	VAERRRQAGIA	240
QY	241	GHTYLOASEKFKI	WGAESAPAPERGYKTGAPGAMDT	SPAASVPAPQVAVRARQV	ASSSSS	300
DB	241	GHTYLOASEKFKI	WGAESAPAPERGYKTGAPGAMDT	SPAASVPAPQVAVRARQV	ASSSSS	300
QY	301	SSFPAPADMNPVS	ASGAPAPPPPGDGSYLWIPASHYNQL	VTGQSA	PRHPPLTACGL	PAAG 360
DB	301	SSFPAPADMNPVS	ASGAPAPPPPGDGSYLWIPASHYNQL	VTGQSA	PRHPPLTACGL	PAAG 360
QY	361	TVAYGHGAGSPHY	PPPPHAPYPCMLFAGSPLEAQIAALVGA	TAADRQAGGL	PAAGD 420	
DB	361	TVAYGHGAGSPHY	PPPPHAPYPCMLFAGSPLEAQIAALVGA	TAADRQAGGL	PAAGD 420	
QY	421	HGIRSAKRRRHEV	QPEYDCGRDPRD	FPYYPGEARPEPRP	VDSRAARQASGPHETI	480
DB	421	HGIRSAKRRRHEV	QPEYDCGRDPRD	FPYYPGEARPEPRP	VDSRAARQASGPHETI	480
QY	481	TALVGAVTSLQ	QELAHMRARTHAPYGPYPVGYHHPHAD	TETPAQPPRYPAKAVYL	PPPP 540	
DB	481	TALVGAVTSLQ	QELAHMRARTHAPYGPYPVGYHHPHAD	TETPAQPPRYPAKAVYL	PPPP 540	
QY	541	HIAPGPPPLSG	AVPPSPYPVAVTPGAPPLHQPSPAH	AHPPPPPGPTPPPAASL	POQE 600	
DB	541	HIAPGPPPLSG	AVPPSPYPVAVTPGAPPLHQPSPAH	AHPPPPPGPTPPPAASL	POQE 600	
QY	601	APGAEAGALVN	ASSAAHVNDTARAADLFV	SQMMGSR 637		
DB	601	APGAEAGALVN	ASSAAHVNDTARAADLFV	SQMMGSR 637		
RESULT 2						
ADJ77800						
ID	ADJ77800	standard;	protein;	637 AA.		
AC	ADJ77800;					
XX	06-MAY-2004	(first entry)				
DT	06-MAY-2004	(first entry)				
DE	Herpes simplex virus 2 UL7 protein.					
KW	virucide; gene therapy; antigen; skin; lymphocyte;					
KW	peripheral blood mononuclear cell;					
KW	cutaneous lymphocyte-associated antigen; HSV infection.					
OS	Human herpesvirus 2.					
PN	WO2004009021-A2.					
XX	29-JAN-2004.					
PF	18-JUL-2003; 2003WO-US022560.					
PR	18-JUL-2002; 2002US-0396791P.					

PA	(UNIW )	UNIV WASHINGTON.				
XX	Koelle DM,	Liu Z, Corey L;				
XX	WPI;	2004-123280/12.				
DR	Identifying an active antigen of a virus that attacks skin for e.g.					
XX	treating herpes simplex virus, comprises isolating lymphocytes from					
PT	peripheral blood mononuclear cells that express cutaneous lymphocyte-					
PT	associated antigen.					
XX	Claim 9;	SEQ ID NO 9;	123pp;	English.		
XX	The invention relates to a method of identifying an immunologically					
CC	active antigen of a virus that attacks skin comprises: (1) isolating					
CC	lymphocytes from peripheral blood mononuclear cells (PBMC) that express					
CC	cutaneous lymphocyte-associated antigen (CLA); (2) identifying a CLA-					
CC	positive lymphocyte that selectively kills cells infected with the virus;					
CC	and (3) determining the identity of the antigen present in the lymphocyte					
CC	identified in (2). The method is useful for identifying an					
CC	immunologically active antigen of a virus that attacks skin for preparing					
CC	a composition for treating or preventing an HSV infection. This sequence					
CC	corresponds to the UL7 protein from Herpes simplex virus type 2.					
XX	Sequence 637 AA;					
SQ	Query Match	100.0%;	Score 3385;	DB 8;	Length 637;	
	Best Local Similarity	100.0%;	Pred. No. 4e-209;			
	Matches 637;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
QY	1	MASAEVRERLEAPL	PDRAVPIYVAGFLALYDSGDPGELALD	PDPTVRAALPPENPL	PI	NDV 60
DB	1	MASAEVRERLEAPL	PDRAVPIYVAGFLALYDSGDPGELALD	PDPTVRAALPPENPL	PI	NDV 60
QY	61	HRARCEVRGLAVVND	PRGPFVGLIACVQLSERVLETAASAAI	FERRGPALSRERLLYL	120	
DB	61	HRARCEVRGLAVVND	PRGPFVGLIACVQLSERVLETAASAAI	FERRGPALSRERLLYL	120	
QY	121	ITNYLPSVLSLTKRGDEVP	DRTLFAHVALCAIGRRLG	TIYVDTSLDAAI	APRHLDP	180
DB	121	ITNYLPSVLSLTKRGDEVP	DRTLFAHVALCAIGRRLG	TIYVDTSLDAAI	APRHLDP	180
QY	181	ATREGVRRERAAEAL	LAGRTWAPGVEAL	THLLSTAVNNMLRDRWSI	VAERRRQAGIA	240
DB	181	ATREGVRRERAAEAL	LAGRTWAPGVEAL	THLLSTAVNNMLRDRWSI	VAERRRQAGIA	240
QY	241	GHTYLOASEKFKI	WGAESAPAPERGYKTGAPGAMDT	SPAASVPAPQVAVRARQV	ASSSSS	300
DB	241	GHTYLOASEKFKI	WGAESAPAPERGYKTGAPGAMDT	SPAASVPAPQVAVRARQV	ASSSSS	300
QY	301	SSFPAPADMNPVS	ASGAPAPPPPGDGSYLWIPASHYNQL	VTGQSA	PRHPPLTACGL	PAAG 360
DB	301	SSFPAPADMNPVS	ASGAPAPPPPGDGSYLWIPASHYNQL	VTGQSA	PRHPPLTACGL	PAAG 360
QY	361	TVAYGHGAGSPHY	PPPPHAPYPCMLFAGSPLEAQIAALVGA	TAADRQAGGL	PAAGD 420	
DB	361	TVAYGHGAGSPHY	PPPPHAPYPCMLFAGSPLEAQIAALVGA	TAADRQAGGL	PAAGD 420	
QY	421	HGIRSAKRRRHEV	QPEYDCGRDPRD	FPYYPGEARPEPRP	VDSRAARQASGPHETI	480
DB	421	HGIRSAKRRRHEV	QPEYDCGRDPRD	FPYYPGEARPEPRP	VDSRAARQASGPHETI	480
QY	481	TALVGAVTSLQ	QELAHMRARTHAPYGPYPVGYHHPHAD	TETPAQPPRYPAKAVYL	PPPP 540	
DB	481	TALVGAVTSLQ	QELAHMRARTHAPYGPYPVGYHHPHAD	TETPAQPPRYPAKAVYL	PPPP 540	
QY	541	HIAPGPPPLSG	AVPPSPYPVAVTPGAPPLHQPSPAH	AHPPPPPGPTPPPAASL	POQE 600	
DB	541	HIAPGPPPLSG	AVPPSPYPVAVTPGAPPLHQPSPAH	AHPPPPPGPTPPPAASL	POQE 600	
QY	601	APGAEAGALVN	ASSAAHVNDTARAADLFV	SQMMGSR 637		
DB	601	APGAEAGALVN	ASSAAHVNDTARAADLFV	SQMMGSR 637		

RESULT 3	AAW71017	ID	AAW71017 standard; protein; 638 AA.
XX	AAW71017	AC	AAW71017
XX	AAW71017	AC	AAW71017
XX	25-MAR-2003 (revised)	DT	25-MAR-2003 (revised)
DT	02-OCT-1995 (first entry)	DT	02-OCT-1995 (first entry)
XX	HSV-2 protease.	DE	HSV-2 protease.
XX	Herpes simplex virus; HSV-2; capsid; UL26 gene.	KW	Herpes simplex virus; HSV-2; capsid; UL26 gene.
OS	Herpes simplex virus type 2.	OS	Herpes simplex virus type 2.
XX	W09506055-A1.	PN	W09506055-A1.
XX	02-MAR-1995.	PD	02-MAR-1995.
XX	19-AUG-1994; 94WO-US009303.	PF	19-AUG-1994; 94WO-US009303.
XX	20-AUG-1993; 93US-00110522.	PR	20-AUG-1993; 93US-00110522.
PR	23-JUN-1994; 94US-00264537.	PR	23-JUN-1994; 94US-00264537.
XX	(SMIK ) SMITHKLINE BEECHAM CORP.	PA	(SMIK ) SMITHKLINE BEECHAM CORP.
XX	Dilella AG, Debouck CW;	PI	Dilella AG, Debouck CW;
XX	WPI; 1995-106803/14.	DR	WPI; 1995-106803/14.
DR	N-PSDB; AAQ84671.	DR	N-PSDB; AAQ84671.
XX	New herpes simplex virus (HSV)-2 protease and capsid protein - used to develop prods. for use in the diagnosis and treatment of HSV-2 infections.	PT	New herpes simplex virus (HSV)-2 protease and capsid protein - used to develop prods. for use in the diagnosis and treatment of HSV-2 infections.
XX	Claim 1; Fig 1; 51pp; English.	PS	Claim 1; Fig 1; 51pp; English.
XX	The sequence is the product of the herpes simplex virus type 2 gene UL26, which is the HSV-2 protease. The protein can be used in the diagnosis and treatment of HSV-2 infections. See also AAW71018-31. (Updated on 25-MAR-2003 to correct FN field.)	CC	The sequence is the product of the herpes simplex virus type 2 gene UL26, which is the HSV-2 protease. The protein can be used in the diagnosis and treatment of HSV-2 infections. See also AAW71018-31. (Updated on 25-MAR-2003 to correct FN field.)
XX	Sequence 638 AA;	SQ	Sequence 638 AA;
Query Match	99.6%;	Score	3370.5; DB 2; Length 638;
Best Local Similarity	99.7%;	Pred. No.	3.4e-208;
Matches	636; Conservative	1; Mismatches	0; Indels
1; Gaps	1;		
Qy	1	MASAEKRLERLEAPLDRAPVIYVAGFLALYDSGDPGELALDPDTVRAALPPENPLPINVD	60
Db	1	MASAEKRLERLEAPLDRAPVIYVAGFLALYDSGDPGELALDPDTVRAALPPENPLPINVD	60
Qy	61	HRARCEGVRVLAVVNDPRGPFVGLIACVOLERVLETAASAAIFERRGPALSREERLLYL	120
Db	61	HRARCEGVRVLAVVNDPRGPFVGLIACVOLERVLETAASAAIFERRGPALSREERLLYL	120
Qy	121	ITNYLPSVLSLTKRGGVPPDRTLFAHVALCAIGRLGTIVYDTSLSDAIAPRHLDP	180
Db	121	ITNYLPSVLSLTKRGGVPPDRTLFAHVALCAIGRLGTIVYDTSLSDAIAPRHLDP	180
Qy	181	ATREGVRRAAEALAGRTWAPGVLEATHILLSTAVNNMLRDRWSLVAERRRQAGIA	240
Db	181	ATREGVRRAAEALAGRTWAPGVLEATHILLSTAVNNMLRDRWSLVAERRRQAGIA	240
Qy	241	GHTYLAQSEKFKIWAESAPAPERYKTKGPGAMDTSPAASVPAPQAVRARQVA-SSSS	299
Db	241	GHTYLAQSEKFKIWAESAPAPERYKTKGPGAMDTSPAASVPAPQAVRARQVA-SSSS	300
Qy	300	SSSFPADPNVPSAGAPAPPPPGDGSYLWIPASHYNQLVGTGQAPRHPPLTACGLPAA	359
Db	301	SSSFPADPNVPSAGAPAPPPPGDGSYLWIPASHYNQLVGTGQAPRHPPLTACGLPAA	360

Matches 634; Conservative 1; Mismatches 2; Indels 1; Gaps 1;  
QY 1 MASAEEMERLEAPLPDRAVPIYVAGFLALYDSGDPGELALDPTVTRAAALPPENPLPINVD 60  
DB 5 MASAEEMERLEAPLPDRAVPIYVAGFLALYDSGDPGELALDPTVTRAAALPPENPLPINVD 64  
QY 61 HRAECVGRVLA VNDPRGPFVGLIACVQLSERVLETAASAAI FERRGPALSREERLLYL 120  
DB 65 HRAECVGRVLA VNDPRGPFVGLIACVQLSERVLETAASAAI FERRGPALSREERLLYL 124  
QY 121 ITNYLPSVLSLTKRGDEVPDRTLFAHVALCAIGRRLGTIVTYDTSDDAAIAPFRHLDP 180  
DB 125 ITNYLPSVLSLTKRGDEVPDRTLFAHVALCAIGRRLGTIVTYDTSDDAAIAPFRHLDP 184  
QY 181 ATREGVREAAEAEALAGRTWAPGVEALHTLLSTAVNNMLRDRWSLVAERRRQAGIA 240  
DB 185 ATREGVREAAEAEALAGRTWAPGVEALHTLLSTAVNNMLRDRWSLVAERRRQAGIA 244  
QY 241 GHTYLOASEKFKIWGAESAPAPERGYKTGAPGAMDTSPAASVPAQVAVRABOVA-SSSS 299  
DB 245 GHTYLOASEKFKIWGAESAPAPERGYKTGAPGAMDTSPAASVPAQVAVRABOVA-SSSS 304  
QY 300 SSSFPAPADMNPVSASGAPAPPPGDSYIWPASHYNQLVTGQAPRHPPLTACGLPAA 359  
DB 305 SSSFPAPADMNPVSASGAPAPPPGDSYIWPASHYNQLVTGQAPRHPPLTACGLPAA 364  
QY 360 GTVAYGHGAGSPSPHYPPPPPAHPYPCMLPAGSPLEAQIAALVGAIAADRQAGGLPAAAG 419  
DB 365 GTVAYGHGAGSPSPHYPPPPPAHPYPCMLPAGSPLEAQIAALVGAIAADRQAGGLPAAAG 424  
QY 420 DHGIRGSAKRRRHEVEQPEYDCGRDEPDRDFPYYPGCEARPEPRVDSRRAARQASGPHE 479  
DB 425 DHGIRGSAKRRRHEVEQPEYDCGRDEPDRDFPYYPGCEARPEPRVDSRRAARQASGPHE 484  
QY 480 ITALVGAVTSLQOELAHMRARTHAPYGPYPVGYHHPHADTETPAQPPRYPAKAVYLP 539  
DB 485 ITALVGAVTSLQOELAHMRARTHAPYGPYPVGYHHPHADTETPAQPPRYPAKAVYLP 544  
QY 540 PHIAPEGPLSGAVPPSPYPVAVTGPAPPLHQSPSAHAHPPPPPPGTTPPPAASLPQP 599  
DB 545 PHIAPEGPLSGAVPPSPYPVAVTGPAPPLHQSPSAHAHPPPPPPGTTPPPAASLPQP 604  
QY 600 EAPGAEAGALVNASSAAHVNDTARAADLFVSOQMGSR 637  
DB 605 EAPGAEAGALVNASSAAHVNDTARAADLFVSOQMGSR 642

RESULT 5

AAR64767  
ID AAR64767 standard; protein; 636 AA.

AC AAR64767;

XX 25-MAR-2003 (revised)

DT 19-JUL-1995 (first entry)

XX HSV-2 protease, ICP35.

XX Protease; HSV-2; protease-inhibitor; virucide; ICP35.

XX Herpes simplex virus type 2.

XX WO9429456-A2.

XX 22-DEC-1994.

XX 25-MAY-1994; 94WO-US005920.

XX 08-JUN-1993; 93US-00073819.

XX 23-MAY-1994; 94US-00245390.

XX (ABBO ) ABBOTT LAB.

PI Steffy KR, Kati WM, Katz L, Mgonigal TP, Sarthy AV, Schoen SE;  
XX WPI; 1995-036483/05.  
DR N-PSDB; AAR6476252.  
XX New Herpes Simplex Virus type 2 protease - used in screening methods for  
PT identifying potential herpes viral protease inhibitor cpds.  
XX Claim 4; Fig 1A-1C; 56pp; English.  
XX HSV-2 DNA (given in AAR6476252) was isolated from Vero cells infected with  
CC HSV-2 strain Ci. The translated amino acid sequence is given in AAR64767.  
CC The DNA was digested with BamHI, sepd. by gel electrophoresis and  
CC transferred to nitrocellulose. A nick-translated probe deriv. from HSV-1  
CC protease was used for hybridization to identify DNA (AAR6476261) encoding  
CC HSV-2 protease (AAR64768) and the protease substrate ICP35 (AAR64769).  
CC (Updated on 25-MAR-2003 to correct PN field.)  
XX Sequence 636 AA;

Query Match 99.0%; Score 3351.5; DB 2; Length 636;  
Best Local Similarity 99.4%; Pred. No. 5.6e-207;  
Matches 633; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

QY 1 MASAEEMERLEAPLPDRAVPIYVAGFLALYDSGDPGELALDPTVTRAAALPPENPLPINVD 60  
DB 1 MASAEEMERLEAPLPDRAVPIYVAGFLALYDSGDPGELALDPTVTRAAALPPENPLPINVD 60  
QY 61 HRAECVGRVLA VNDPRGPFVGLIACVQLSERVLETAASAAI FERRGPALSREERLLYL 120  
DB 61 HRAECVGRVLA VNDPRGPFVGLIACVQLSERVLETAASAAI FERRGPALSREERLLYL 120  
QY 121 ITNYLPSVLSLTKRGDEVPDRTLFAHVALCAIGRRLGTIVTYDTSDDAAIAPFRHLDP 180  
DB 121 ITNYLPSVLSLTKRGDEVPDRTLFAHVALCAIGRRLGTIVTYDTSDDAAIAPFRHLDP 180  
QY 181 ATREGVREAAEAEALAGRTWAPGVEALHTLLSTAVNNMLRDRWSLVAERRRQAGIA 240  
DB 181 ATREGVREAAEAEALAGRTWAPGVEALHTLLSTAVNNMLRDRWSLVAERRRQAGIA 240  
QY 241 GHTYLOASEKFKIWGAESAPAPERGYKTGAPGAMDTSPAASVPAQVAVRABOVA-SSSS 300  
DB 241 GHTYLOASEKFKIWGAESAPAPERGYKTGAPGAMDTSPAASVPAQVAVRABOVA-SSSS 299  
QY 301 SSSFPAPADMNPVSASGAPAPPPGDSYIWPASHYNQLVTGQAPRHPPLTACGLPAA 360  
DB 300 SSSFPAPADMNPVSASGAPAPPPGDSYIWPASHYNQLVTGQAPRHPPLTACGLPAA 359  
QY 361 TVAYGHGAGSPSPHYPPPPPAHPYPCMLPAGSPLEAQIAALVGAIAADRQAGGLPAAAG 420  
DB 360 TVAYGHGAGSPSPHYPPPPPAHPYPCMLPAGSPLEAQIAALVGAIAADRQAGGLPAAAG 419  
QY 421 HGIRGSAKRRRHEVEQPEYDCGRDEPDRDFPYYPGCEARPEPRVDSRRAARQASGPHE 480  
DB 420 HGIRGSAKRRRHEVEQPEYDCGRDEPDRDFPYYPGCEARPEPRVDSRRAARQASGPHE 479  
QY 481 TALVGAVTSLQOELAHMRARTHAPYGPYPVGYHHPHADTETPAQPPRYPAKAVYLP 540  
DB 480 TALVGAVTSLQOELAHMRARTHAPYGPYPVGYHHPHADTETPAQPPRYPAKAVYLP 539  
QY 541 HIAPEGPLSGAVPPSPYPVAVTGPAPPLHQSPSAHAHPPPPPPGTTPPPAASLPQP 600  
DB 540 HIAPEGPLSGAVPPSPYPVAVTGPAPPLHQSPSAHAHPPPPPPGTTPPPAASLPQP 599  
QY 601 AFGAEAGALVNASSAAHVNDTARAADLFVSOQMGSR 637  
DB 600 AFGAEAGALVNASSAAHVNDTARAADLFVSOQMGSR 636

RESULT 6

AAR64767  
ID AAR64767 standard; protein; 657 AA.



AC	AAW72192;	QY	360	GTVAYGHPGAGSPHYPPPPPAHPYPGMLFAGPSPLEAQIAALVGAIAADRQAGLPAAAG	419
XX		DB	361	GTVAYGHPGAGSPHYPPPPPAHPYPGMLFAGPSPLEAQIAALVGAIAADRQAGLPAAAG	420
DE	HSV-2 strain SB5 Contig ID 15 ORF#27 protein.	QY	420	DHGIRGSAKRREHEVEQPEYDCGRDEPDRDPFYYPGEARPEPRPVDSSRAARQASGPHET	479
XX		DB	421	DHGIRGSAKRREHEVEQPEYDCGRDEPDRDPFYYPGEARPEPRPVDSSRAARQASGPHET	480
KW	HSV-2 strain SB5; immunological response induction; therapy;	QY	480	ITALVGAVTSLQQLAHMRARTHAPYGPYPVGVGYHHPHADTETPAQPPRYPAKAVYLP	539
XX	antiviral identification; viral protein inhibitor.	DB	481	ITALVGAVTSLQQLAHMRARTHAPYGPYPVGVGYHHPHADTETPAQPPRYPAKAVYLP	540
OS	Herpes simplex virus 2.	QY	540	PHIAPPGLPLSGAVPPPSYPVAVTGPAPPLHQPSPAHAPHPPPPPPPPPPPPPPPPP	599
XX	WO9820016-A1.	DB	541	PHIAPPGLPLSGAVPPPSYPVAVTGPAPPLHQPSPAHAPHPPPPPPPPPPPPPPPPP	600
XX	14-MAY-1998.	QY	600	EAPGAENGALVNASSAAHV 618	
XX	31-OCT-1997; 97WO-US020016.	DB	601	EAPGAENGALVNASSAAHV 619	
PR	04-NOV-1996; 96US-0030279P.				
PR	09-JUN-1997; 97US-0049018P.				
XX	(SMIK ) SMITHKLINE BEECHAM CORP.				
PI	Essex KM, Chan JY, Dabrowski-Amara CE, Delvecchio AM, Dillon SB;				
PI	Leary JJ;				
XX	WPI; 1998-286847/25.				
DR	N-PSDB; AAV62176.				
XX					
XX	Herpes simplex virus type-2 sequences - useful in, e.g. prevention and				
PT	treatment of infection or inducing immunological response in mammal.				
XX	Claim 10; Page 120; 748pp; English.				
XX	This sequence represents a Herpes simplex virus type-2 (HSV-2) protein				
CC	sequence of the invention. This sequence was isolated from a HSV-2 strain				
CC	SB5 (deposited as ARCC VR-2546) DNA fragment designated Contig ID 15.				
CC	Based on homology, this sequence is a capsid protein P40. The proteins				
CC	can be used for the treatment or prevention of disease, to induce an				
CC	immunological response in a mammal or to identify inhibitors, activators				
CC	or novel antivirals. Antagonists of the proteins can be used to inhibit a				
CC	viral polypeptide. The DNA sequence or a vector containing it can also be				
CC	used to induce an immunological response in a mammal				
XX					
SQ	Sequence 657 AA;				
	Query Match 96.5%; Score 3267.5; DB 2; Length 657;				
	Best Local Similarity 99.4%; Pred. No. 1.4e-201;				
	Matches 615; Conservative 1; Mismatches 2; Indels 1; Gaps 1;				
QY	1 MASAEMRERLEAPLPDRAVPIYVAGFLALYDSGDPGELALDPDTVRAALPPENPLPINVD	60			
DB	1 MASAEMRERLEAPLPDRAVPIYVAGFLALYDSGDPGELALDPDTVRAALPPENPLPINVD	60			
QY	61 HRARCEVGRVLAVNDPRGPFVGLIACVOLERVLETAASAAIFERRGPALSREERLLYL	120			
DB	61 HRARCEVGRVLAVNDPRGPFVGLIACVOLERVLETAASAAIFERRGPALSREERLLYL	120			
QY	121 ITNVLPSVLSLTKRGDEVPDRTLFAHVALCAIGRLGTIVTYDTSLDAAIAPFRHLDP	180			
DB	121 ITNVLPSVLSLTKRGDEVPDRTLFAHVALCAIGRLGTIVTYDTSLDAAIAPFRHLDP	180			
QY	181 ATREGVRRERAAEALAGTWPAGVEALHTLLSTAVNNMLRDRWSLVAERRRQAGIA	240			
DB	181 ATREGVRRERAAEALAGTWPAGVEALHTLLSTAVNNMLRDRWSLVAERRRQAGIA	240			
QY	241 GHTYLQASEFKIWGASAPAPRGYKGTGAPGAMDTSPAASVPAPQVAVRQVA-SSSS	299			
DB	241 GHTYLQASEFKIWGASAPAPRGYKGTGAPGAMDTSPAASVPAPQVAVRQVA-SSSS	300			
QY	300 SSSFPAPADNPNVSASGAPAPPPGDSYLIWIPASHYNQLVTGQSAPRHPPLTACGLPAA	359			
DB	301 SSSFPAPADNPNVSASGAPAPPPGDSYLIWIPAFHYNQLVTGQSAPRHPPLTACGLPAA	360			
	Query Match 68.9%; Score 2333; DB 2; Length 635;				
	Best Local Similarity 74.0%; Pred. No. 1.4e-141;				
	Matches 477; Conservative 30; Mismatches 120; Indels 18; Gaps 7;				
QY	1 MASAEMRERLEAPLPDRAVPIYVAGFLALYDSGDPGELALDPDTVRAALPPENPLPINVD	60			
DB	1 MASAEMRERLEAPLPDRAVPIYVAGFLALYDSGDPGELALDPDTVRAALPPENPLPINVD	60			

RESULT 7  
AAR71031  
ID AAR71031 standard; protein; 635 AA.  
XX  
AC AAR71031;  
XX  
AC AAR71031;  
XX  
DT 25-MAR-2003 (revised)  
DT 02-OCT-1995 (first entry)  
XX  
DE HSV-1 UL26 gene product.  
XX  
KW Herpes simplex virus; HSV-1 protease; capsid.  
XX  
OS Herpes simplex virus type 1.  
XX  
PN WO9506055-A1.  
XX  
PD 02-MAR-1995.  
XX  
PF 19-AUG-1994; 94WO-US009303.  
XX  
PR 20-AUG-1993; 93US-00110522.  
PR 23-JUN-1994; 94US-00264537.  
XX  
XX (SMIK ) SMITHKLINE BEECHAM CORP.  
XX  
XX Dilella AG, Debouck CM;  
XX  
XX WPI; 1995-106803/14.  
XX  
XX N-PSDB; AAQ84678.  
XX  
XX New herpes simplex virus (HSV)-2 protease and capsid protein - used to  
XX develop prods. for use in the diagnosis and treatment of HSV-2  
XX infections.  
XX  
XX Disclosure; Page 35; 51pp; English.  
XX  
XX The sequence is that of a herpes simplex virus type 1 protease and  
XX capsid. The protein can be used in the diagnosis and treatment of HSV-1  
XX infections. See also R717017-30. (Updated on 25-MAR-2003 to correct PN  
XX field.)  
XX  
XX Sequence 635 AA;



QY 535 VILPPPHIAPPPLSGA--VPPPSYPPVAVTGPAPPLHQPSPAHAPPPPPPGTPPPP 592  
 DB 534 VYRPPPHSAFYGGPPQSPASHPTPPYAPACPGPPPP---PCPSTQTRAPLPTPEAFPP 590  
 QY 593 AASLPQPEAGAGALVNASAAHVNDTARAADLFVSQMMGSR 637  
 DB 591 AATGSQPEASNAEAGALVNASAAHVNDTARAADLFVSQMMGAR 635

RESULT 9  
 AAR28634  
 ID AAR28634 standard; protein; 635 AA.  
 AC AAR28634;  
 XX  
 DT 25-MAR-2003 (revised)  
 DT 24-MAR-1993 (first entry)  
 DE UL26 protease.  
 XX  
 KW UL26; open reading frame; ORF; protease; herpes simplex virus; HSV;  
 KW capsid protein; ICP35; ICP35 c; ICP35 d; ICP35 e; ICP35 f;  
 KW functional domain.  
 XX  
 OS Herpes simplex virus type 1.  
 XX  
 PN EP514830-A2.  
 XX  
 PD 25-NOV-1992.  
 XX  
 PF 19-MAY-1992; 92EP-00108420.  
 XX  
 PR 24-MAY-1991; 91US-00705814.  
 PR 07-FEB-1992; 92US-00832855.  
 XX  
 PA (ARCH-) ARCH DEV CORP.  
 XX  
 PI Roizman B, Liu F;  
 XX  
 WPI; 1992-391444/48.  
 DR N-PSDB; AAQ30736.  
 XX  
 PT New herpes serine proteases and corresp. nucleic acid sequences - for  
 PT detection, prevention and treatment of infection caused by HSV, Epstein-  
 PT Barr, Varicella-Zoster and CMV cytomegalo-virus.  
 XX  
 PS Disclosure; Fig 1B; 66pp; English.  
 XX  
 CC This sequence is encoded by the UL26 open reading frame (ORF) and it is a  
 CC protease. The UL26 ORF is contained within a section of the herpes  
 CC simplex virus (HSV) type 1 genome for the family of HSV capsid proteins  
 CC designated ICP35. The UL26 transcription initiation site is used as the  
 CC start point for measurements of distance in this section of the genome.  
 CC The UL26 protease and ICP35 precursor proteins, ICP35 c and d, may be  
 CC cleaved by the UL26 protease. Cleavage of ICP35 c and d gives ICP35 e and  
 CC f respectively. Cleavage specificity of the UL26 protease may be altered  
 CC by mutations within the sequence. These mutations may be insertions,  
 CC deletions or substitutions which affect the functional domains of the  
 CC protease. See also AAR28635-52. (Updated on 25-MAR-2003 to correct PN  
 CC field.)  
 XX  
 SQ Sequence 635 AA;

Query Match 58.8%; Score 2329; DB 2; Length 635;  
 Best Local Similarity 73.8%; Pred. No. 2.5e-141;  
 Matches 476; Conservative 30; Mismatches 121; Indels 18; Gaps 7;

QY 1 MASAENRERLEAPDRAPVIYVAGFLALYDSGDPCELALDPTVRAALPPENPLPINVD 60  
 DB 1 MAADAPGRNEEPDRAPVIYVAGFLALYDSGDELALDPTVRAALPPDNPPLPINVD 60  
 QY 61 HRACEVGRVLAVVNDPRGPFVGLIACVQLERVLETAASAAIFERRGPALSREERLLYL 120

DB 61 HRACEVGRVLAVVNDPRGPFVGLIACVQLERVLETAASAAIFERRGPALSREERLLYL 120  
 QY 121 ITNLYSVSLSTKRGDEVPPDRTLFAHVALCAIGRELGTIVTYDTSLDAAIAFPRLDP 180  
 DB 121 ITNLYSVSLATKRLGGEAHPDRTLFAHVALCAIGRRLGTIVTYDGLDAAIAFPRLSP 180  
 QY 181 ATREGVRRERAAEALALAGRTWAPGVEALHTLLSTAVNNMLRDRWSLVAERRRQAGIA 240  
 DB 181 ASREGARRLAEEAELALSGRTWAPGVEALHTLLSTAVNNMLRDRWSLVAERRRQAGIA 240  
 QY 241 GHTYLOASEKFKIWAESAAPAPERGYKTGAPGAMDTSPAASVPAPQ----VAVRAROVAS 296  
 DB 241 GHTYLOASEKFKMGAEPVSAPARGYKNGAPESTDIPPGSIAAAPQGDRCPIVRQGVAL 300  
 QY 297 SSSSSFFPAPADNPNVPSASGAPAPPDGGSGSYLWIPASHYNQLVGTGOSAPRHPLTACGL 356  
 DB 301 S-----PVLPPNVPVTSPTAPAPDGGSGSYLWIPASHYNQLVAGHAAFPQPHSAFGP 354  
 QY 357 P-AAGTVAYGHPGAGSPHYPPPPHAPYPCGMLFAGSPLEAQIAALVGAIAADRQAGLP 415  
 DB 355 PAAAGSVAYGHPGAGLSQHYPPHVAHQYPGVLFSGSPLEAQIAALVGAIAADRQAGQP 414  
 QY 416 AAAGDHGIRGSAKRRHEVEQPEYDCGRDEPDRDPFYYPGCEARPEPRPVDSRAARQASG 475  
 DB 415 -AAGDPGVRGSGKRRRYEAGPSESYCDQDEPDADYPYPGCEARGAPRGVDSRAARHSPG 473  
 QY 476 PHETITLVGAVTSLQOELAHMREARTHAPYGPVPPVPGVPHHHPADTE-TPAQPPRYPAKA 534  
 DB 474 TNETITLVGAVTSLQOELAHMREARTSAPYGMTTPVAHYRPQVGEPEPTTTHALCPDEA 533  
 QY 535 VILPPPHIAPPGPPPLS--GAVPPPSYPPVAVTGPAPPLHQPSPAHAPPPPPPGTPPPP 592  
 DB 534 VYRPPPHSAFYGGPPQSPASHPTPPYAPACPGPPPP---PCPSTQTRAPLPTPEAFPP 590  
 QY 593 AASLPQPEAGAGALVNASAAHVNDTARAADLFVSQMMGSR 637  
 DB 591 AATGSQPEASNAEAGALVNASAAHVNDTARAADLFVSQMMGAR 635

RESULT 10  
 AAR28651  
 ID AAR28651 standard; protein; 635 AA.  
 AC AAR28651;  
 XX  
 DT 25-MAR-2003 (revised)  
 DT 24-MAR-1993 (first entry)  
 DE UL26 protease substitution mutant KK.  
 XX  
 KW UL26; open reading frame; ORF; protease; herpes simplex virus; HSV;  
 KW capsid protein; ICP35; ICP35 c; ICP35 d; ICP35 e; ICP35 f;  
 KW functional domain; insertion; deletion; substitution.  
 XX  
 OS Herpes simplex virus type 1.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 215  
 FT /label= S215A  
 XX  
 PN EP514830-A2.  
 XX  
 PD 25-NOV-1992.  
 XX  
 PF 19-MAY-1992; 92EP-00108420.  
 XX  
 PR 24-MAY-1991; 91US-00705814.  
 PR 07-FEB-1992; 92US-00832855.  
 XX  
 PA (ARCH-) ARCH DEV CORP.  
 XX  
 PI Roizman B, Liu F;



Db 181 ASREGARRLAABAEALSGRTWARGVEALHTLLSTAVNNMLDRWSLVAERRRQAGIA 240  
 Qy 241 GHTYLOASEFKIWAESAEPAPERGKYGAPGAMDTSPAASVPAPQ----VAVRAROVAS 296  
 Db 241 GHTYLOASEFKIWAESAEPVAPARGYKNGAPESTDIPPGSTAAAPQGDRCPIVQRGVAL 300  
 Qy 297 SSSSSFPADMPNVSAGAPAPPPGDSYLMIPASHYNQLVTGOSAPRHPPLTACGL 356  
 Db 301 S-----PVLPMNPVPTSGTPAPPGDSYLMIPASHYNQLVAGHAAPQOPHSAFGF 354  
 Qy 357 P-AAGTVAYGHPGAGSPHPYPPPAHPVPGMLPAGSPLEAQIAALVGAIAADRQAGGLP 415  
 Db 355 PAAGSVAYGHPGAGLSQHTPPHVAHQVPGVLFSPGLEAQIAALVGAIAADRQAGQP 414  
 Qy 416 AAAGDHGIRGSAKRREHEVQPEYDCGRDEPRDFPYYPGEARPEPRVDSRRAARQASG 475  
 Db 415 -AAGDPGVRGSKRRRYEAGPSESICYDQDEPDADYPYPGEARGAPRGVDSRRAARHSPG 473  
 Qy 476 PHETITLVGAVTSLOQELAHMRARTHAPYGPYPVPGYHHPHADTETP--AOPPRYPAP 533  
 Db 474 TNETITLVGAVTSLOQELAHMRARTSAPYGMVTFVAYRXPQVGEPEPTTTHPALCPPE 533  
 Qy 534 AVYLPPIHAPPGPPLS--GAVPPSPYPVAVTPGAPPLHQPSPAHAPHPPPPPGPTTP 591  
 Db 534 AVYRPPHSAFYGPPQGPSPHAPTPPYAPAAACPPGPPPP---PCPSTQTRAPLPTPEAFP 590  
 Qy 592 PAASLOPEAPGAEAGALVNASSAAHVNVDTRAADLFSVQMNGSR 637  
 Db 591 PAATGSQPEASNAEAGALVNASSAAHVNVDTRAADLFSVQMNGAR 636

## RESULT 12

AAR28647  
 ID AAR28647 standard; protein; 635 AA.

AC AAR28647;  
 XX

DT 25-MAR-2003 (revised)

DT 24-MAR-1993 (first entry)

DE UL26 protease substitution mutant GG.

XX UL26; open reading frame; ORF; protease; herpes simplex virus; HSV;  
 KW capsid protein; ICP35; ICP35 c; ICP35 d; ICP35 e; ICP35 f;  
 KW functional domain; insertion; deletion; substitution.

OS Herpes simplex virus type 1.

XX Key Location/Qualifiers  
 FH Misc-difference 7. .9  
 FT /note= "Substitution GDR > SRT"

XX EF514830-A2.

XX 25-NOV-1992.

XX 19-MAY-1992; 92EP-00108420.

XX 24-MAY-1991; 91US-00705814.

XX 07-FEB-1992; 92US-00832855.

XX (ARCH-) ARCH DEV CORP.

XX Roizman B, Liu F;

XX WPI; 1992-39144/48.

XX New herpes serine proteases and corresp. nucleic acid sequences - for

XX detection, prevention and treatment of infection caused by HSV, Epstein-

XX Barr, Varicella-Zoster and CMV cytomegalo-virus.

XX Disclosure; Table 1; 66pp; English.

XX

CC The sequences given in AAR28635-52 are mutants of the protease encoded by  
 CC the UL26 open reading frame (ORF). The UL26 ORF is contained within a  
 CC section of the herpes simplex virus (HSV) type 1 genome for the family of  
 CC HSV capsid proteins designated ICP35. The UL26 transcription initiation  
 CC site is used as the start point for measurements of distance in this  
 CC section of the genome. The UL26 protease and ICP35 precursor proteins,  
 CC ICP35 c and d, may be cleaved by the UL26 protease. Cleavage of ICP35 c  
 CC and d gives ICP35 e and f respectively. Cleavage specificity of the UL26  
 CC protease is altered by the inclusion of these mutations. These mutations  
 CC include insertions, deletions or substitutions which affect the functional  
 CC domains of the protease. (Updated on 25-MAR-2003 to correct PN field.)  
 XX  
 SQ Sequence 635 AA;

Query Match 68.6%; Score 2322; DB 2; Length 635;

Best Local Similarity 73.6%; Pred. No. 7e-141;

Matches 475; Conservative 29; Mismatches 123; Indels 18; Gaps 7;

Qy 1 MASAEMRERLEAPLDRAPVIYVAGFLALYDSGDPGLALDDPTVRAALPENPLPINVD 60  
 Db 1 MAADAPSRTMEELPDRAPVIYVAGFLALYDSGDPGLALDDPTVRAALPENPLPINVD 60  
 Qy 61 HRARCEVGRVLAVNDPRGPFVGLIACVOLERVLETAASAAIFEREGPALSRERLLYL 120  
 Db 61 HRAGCEVGRVLAVNDPRGPFVGLIACVOLERVLETAASAAIFEREGPALSRERLLYL 120  
 Qy 121 ITNVLPSVLSLTKRRGDEVDPDRTLFAHVALCAIGRLGTIVTYDTSIDAIAAPFRHLDP 180  
 Db 121 ITNVLPSVLSLTKRLGGEHPDRTLFAHVALCAIGRLGTIVTYDTSIDAIAAPFRHLSP 180  
 Qy 181 ATREGVREAAEAELALAGRTWAPGVEALHTLLSTAVNNMLDRWSLVAERRRQAGIA 240  
 Db 181 ASREGARRLAABAEALSGRTWAPGVEALHTLLSTAVNNMLDRWSLVAERRRQAGIA 240  
 Qy 241 GHTYLOASEFKIWAESAEPAPERGKYGAPGAMDTSPAASVPAPQ----VAVRAROVAS 296  
 Db 241 GHTYLOASEFKIWAESAEPVAPARGYKNGAPESTDIPPGSTAAAPQGDRCPIVQRGVAL 300  
 Qy 297 SSSSSFPADMPNVSAGAPAPPPGDSYLMIPASHYNQLVTGOSAPRHPPLTACGL 356  
 Db 301 S-----PVLPMNPVPTSGTPAPPGDSYLMIPASHYNQLVAGHAAPQOPHSAFGP 354  
 Qy 357 P-AAGTVAYGHPGAGSPHPYPPPAHPVPGMLPAGSPLEAQIAALVGAIAADRQAGGLP 415  
 Db 355 PAAGSVAYGHPGAGLSQHTPPHVAHQVPGVLFSPGLEAQIAALVGAIAADRQAGQP 414  
 Qy 416 AAAGDHGIRGSAKRREHEVQPEYDCGRDEPRDFPYYPGEARPEPRVDSRRAARQASG 475  
 Db 415 -AAGDPGVRGSKRRRYEAGPSESICYDQDEPDADYPYPGEARGAPRGVDSRRAARHSPG 473  
 Qy 476 PHETITLVGAVTSLOQELAHMRARTHAPYGPYPVPGYHHPHADTETP--AOPPRYPAP 534  
 Db 474 TNETITLVGAVTSLOQELAHMRARTSAPYGMVTFVAYRXPQVGEPEPTTTHPALCPPEA 533  
 Qy 535 VYLPPIHAPPGPPLS--GAVPPSPYPVAVTPGAPPLHQPSPAHAPHPPPPPGPTTPP 592  
 Db 534 VYRPPHSAFYGPPQGPSPHAPTPPYAPAAACPPGPPPP---PCPSTQTRAPLPTPEAFP 590  
 Qy 593 AASLPOPEAPGAEAGALVNASSAAHVNVDTRAADLFSVQMNGSR 637  
 Db 591 AATGSQPEASNAEAGALVNASSAAHVNVDTRAADLFSVQMNGAR 635

## RESULT 13

AAR28652

ID AAR28652 standard; protein; 635 AA.

XX AAR28652;

XX AC AAR28652;

XX DT 25-MAR-2003 (revised)

XX DT 24-MAR-1993 (first entry)

XX UL26 protease substitution mutant LL.

DE



Query Match		68.5%; Score 2319; DB 2; Length 626;
Best Local Similarity		74.4%; Pred. No. 1.1e-140;
Matches		473; Conservative 26; Mismatches 117; Indels 18; Gaps 7;
QY	10	LEAPLPDRAPIYVAGFLALYDSDGPGELALDPTVRAALPPENPLPINDVHRARCEVGR 69
DB	1	MEEPPLDRAPIYVAGFLALYDSDGPGELALDPTVRAALPPENPLPINDVHRARCEVGR 60
QY	70	VLAIVNDPRGPFVGLIACVQLSERVLETAASAAIFERRGPALGREERLLYLITNYPVS 129
DB	61	VLAIVDDPRGPFVGLIACVQLSERVLETAASAAIFERRGPALGREERLLYLITNYPVS 120
QY	130	LSTKRGEVDPDRTLFAHVALCAIGRLGTIVTYDTSLSAAIAPFPHLDPATREGVRE 189
DB	121	LATKRLGGEAHPDRTLFAHVALCAIGRLGTIVTYDTSLSAAIAPFPHLDPATREGVRE 180
QY	190	AAAEALALAGRTWAPGVEALTHTLSTAVNNMLDRWSLVAERRRQAGTACHTYLOASE 249
DB	181	AAAEALALAGRTWAPGVEALTHTLSTAVNNMLDRWSLVAERRRQAGTACHTYLOASE 240
QY	250	KFKIWGAESAPAPRGYKTCAPGAMDTSPAASVPAPQ- ---VAVRAQVASSSSSSSFFPA 305
DB	241	KFKIWGAEPVSAARGYKNGAPESTDIPGSIAPAPQDCRCPIVRQGVALS- ---PV 294
QY	306	PADWNPVSAAGAPAPPDGGSYLWIPASHYNQLVTQSAPRHPPLTACGLP- AAGTVAY 364
DB	295	LPPWNPVPTSGTPAPAPPDGGSYLWIPASHYNQLVAGHAAPQPQPHSAFGFPAAAGSVAY 354
QY	365	CHFGAGSPHYPPPPHAPHPGMLFAGSPLEAQIAALVGAIAADRQAGGLPAAAGDHGIR 424
DB	355	GPHGAGLSQHYPPHVAHQYGVLFSGSPLEAQIAALVGAIAADRQAGGQP- AAGDPGVR 413
QY	425	GSAKRRHEVEQDEYDCGRDEPRDPYPYGEARPEPRPVDSRAARQASGHETITALV 484
DB	414	GSKRRRYEAGPSYCDQDEPDADPIYFEGEARGAPRGVDSRAARHSPGTNETITALM 473
QY	485	GAVTSLQOELAHMARHTHAPYGPYPVGPYHHPHADTE-TPAQPPRYPAKAVYLPHPHIA 543
DB	474	GAVTSLQOELAHMARHTSAPYGMVTPVAHYRPQVGBEPETTTTHPALCPPEAVYRPHPHA 533
QY	544	PPGPPPLS- -GAVPPSPYPPVAVTPGAPPLHQSPAHAPPPPPPGTPPPPAASLPQPEA 601
DB	534	PYGPPQGPSPSHAPTPPYAPACPPGPPPP- -PCPSTQTRAPLPTSPAPFAATGSOPEA 590
QY	602	PGAEAGALVNASSAAHVNDTARAADLFVSQMMGSR 637
DB	591	SNABAGALVNASSAAHVNDTARAADLFVSQMMGAR 626
RESULT 15		
ID	AA28650	
XX	AA28650	standard; protein; 635 AA.
AC	AA28650;	
XX	AA28650;	
DT	25-MAR-2003	(revised)
DT	24-MAR-1993	(first entry)
XX	UL26	protease substitution mutant J7.
XX	UL26;	open reading frame; ORF; protease; herpes simplex virus; HSV;
KW	capsid protein; ICP35;	ICP35 c; ICP35 d; ICP35 e; ICP35 f;
KW	functional domain;	insertion; deletion; substitution.
OS	Herpes simplex virus type 1.	
XX	Key	Location/Qualifiers
FX	FT	Misc-difference 148
FT	XX	/label= H148A
XX	EP514830-A2.	
XX		

25-NOV-1992.

19-MAY-1992; 92EP-00108420.

24-MAY-1991; 91US-00705814.

07-FEB-1992; 92US-00832855.

(ARCH-) ARCH DEV CORP.

Roizman B, Liu F;

WPI; 1992-391444/48.

New herpes serine proteases and corresp. nucleic acid sequences - for

detection, prevention and treatment of infection caused by HSV, Epstein-

Barr, Varicella-Zoster and CMV cytomegalo-virus.

Disclosure; Table 1; 66pp; English.

The sequences given in AAR28635-52 are mutants of the protease encoded by the UL26 open reading frame (ORF). The UL26 ORF is contained within a section of the herpes simplex virus (HSV) type 1 genome for the family of HSV capsid proteins designated ICP35. The UL26 transcription initiation site is used as the start point for measurements of distance in this section of the genome. The UL26 protease and ICP35 precursor proteins, ICP35 c and d, may be cleaved by the UL26 protease. Cleavage of ICP35 c and d gives ICP35 e and f respectively. Cleavage specificity of the UL26 protease is altered by the inclusion of these mutations. These mutations include insertions, deletions or substitutions which affect the functional domains of the protease. (Updated on 25-MAR-2003 to correct PN field.)

Query Match 68.5%; Score 2319; DB 2; Length 635;  
 Best Local Similarity 73.8%; Pred. No. 1.1e-140;  
 Matches 475; Conservative 30; Mismatches 122; Indels 18; Gaps 7;

Sequence 635 AA;

Query Match 68.5%; Score 2319; DB 2; Length 635;

Best Local Similarity 73.8%; Pred. No. 1.1e-140;

Matches 475; Conservative 30; Mismatches 122; Indels 18; Gaps 7;

QY 1 MASAENRERLEAPLPDRAPIYVAGFLALYDSDGPGELALDPTVRAALPPENPLPINDV 60

DB 1 MAADAPGDRMEELPLDRAPIYVAGFLALYDSDGPGELALDPTVRAALPPENPLPINDV 60

QY 61 HRARCEVGRVLAIVNDPRGPFVGLIACVQLSERVLETAASAAIFERRGPALGREERLLYL 120

DB 61 HRAGCEVGRVLAIVDDPRGPFVGLIACVQLSERVLETAASAAIFERRGPALGREERLLYL 120

QY 121 ITNYPVSLSLTKRCGEVDPDRTLFAHVALCAIGRLGTIVTYDTSLSAAIAPFPHLDP 180

DB 121 ITNYPVSLSLTKRCGEVDPDRTLFAHVALCAIGRLGTIVTYDTSLSAAIAPFPHLSP 180

QY 181 ATREGVRRERAAEALALAGRTWAPGVEALTHTLSTAVNNMLDRWSLVAERRRQAGIA 240

DB 181 ASREGARRLAAEALALAGRTWAPGVEALTHTLSTAVNNMLDRWSLVAERRRQAGIA 240

QY 241 GHTYLOASEKFTIWGAESAPAPRGYKTCAPGAMDTSPAASVPAPQ- ---VAVRAQVAS 296

DB 241 GHTYLOASEKFTIWGAEPVSAARGYKNGAPESTDIPGSIAPAPQDCRCPIVRQGV 300

QY 297 SSSSSFFPAPADWNPVSAAGAPAPPDGGSYLWIPASHYNQLVTQSAPRHPPLTACGL 356

DB 301 S-----PVLPPMNPVPTSGTPAPAPPDGGSYLWIPASHYNQLVAGHAAPQPQPHSA 354

QY 357 P-AAGTVAYGHPGAGSPHYPPPPHAPHPGMLFAGSPLEAQIAALVGAIAADRQAGGLP 415

DB 355 PAAGSVAYGHPGAGLSQHYPPHVAHQYGVLFSGSPLEAQIAALVGAIAADRQAGGQP 414

QY 416 AAAGDHGIRGSAKRRRHEVEQDEYDCGRDEPRDPYPYGEARPEPRPVDSRAARQASG 475

DB 415 -AAGDPGVRGSKRRRYEAGPSYCDQDEPDADPIYFEGEARGAPRGVDSRAARHSPG 473

QY 476 PHETITLVCANTSLQOELAHMARHTHAPYGPYPVGPYHHPHADTE-TPAQPPRYPAKA 534

DB 474 TNETITLGVAVTSLQOELAHMARHTSAPYGMVTPVAHYRPQVGBEPETTTTHPALCPPEA 533

